CERES MODIS Cloud Product Retrievals for Edition 4, Part II: Comparisons to CloudSat and CALIPSO

Journal:	*Transactions on Geoscience and Remote Sensing*
Manuscript ID	Draft
Manuscript Type:	Regular paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Yost, Christopher; SSAI, Clouds & Surface Retrievals Minnis, Patrick; SSAI, Remote Sensing; Sun-Mack, Sunny; SSAI, Clouds Retrievals Chen, Yan; Science Systems and Applications Inc Hampton, Climate Science Branch Smith, Bill; NASA LaRC, Climate Sciences Branch
Keywords:	Atmosphere, Multispectral Data, Geophysical Data

CERES MODIS Cloud Product Retrievals for Edition 4, Part II: Comparisons to CloudSat and CALIPSO

Christopher R. Yost, Patrick Minnis, Szedung Sun-Mack, Yan Chen, William L. Smith, Jr.

Abstract— **Assessments of the Clouds and the Earth's Radiant Energy System Edition 4 (Ed4) cloud retrievals are critical for climate studies. Ed4 cloud parameters are evaluated using instruments in the A-Train Constellation. Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) and Cloud Profiling Radar (CPR) retrievals are compared to Ed4 retrievals from the Aqua Moderate-Resolution Imaging Spectroradiometer (MODIS) as a function of CALIOP horizontal averaging (HA) scale. Regardless of HA scale, MODIS daytime (nighttime) water cloud fraction is greater (less) than that from CALIOP. MODIS ice cloud fraction is less than CALIOP overall, with largest differences in polar regions. Ed4 and CALIOP retrieve the same cloud phase in 70-98% of simultaneous observations depending on time of day, surface conditions, HA scales, and type of cloud vertical structure. Mean cloud top height differences for single-layer water clouds over snow/ice-free surfaces are less than 100 m. Base altitude positive biases of 170 – 460 m may be impacted by CPR detection limitations. Average MODIS ice cloud top heights are underestimated by 70 m for some deep convective clouds and up to ~2.2 km for thin cirrus. Ice cloud base altitudes are typically underestimated (overestimated) during daytime (nighttime). MODIS and CALIOP cirrus optical depths over oceans are within 46% and 5% for daytime and nighttime observations, respectively. Ice water path differences depend on the CALIOP retrieval version and warrant further investigation. Except for daytime cirrus optical depth, Ed4 cloud property retrievals are at least as accurate as other long-term operational cloud property retrieval systems.**

***Index Terms*—CALIPSO, climate, cloud, cloud height, cloud optical depth, cloud phase, cloud remote sensing Clouds and the Earth's Radiant Energy System (CERES), MODerate-resolution Imaging Spectroradiometer (MODIS), validation**

NOMENCLATURE

A-Train	Afternoon Satellite Constellation
CA	CALIPSO
CALIOP	Cloud-Aerosol Lidar with Orthogonal Polarization
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder in Space Observations
CERES	Clouds and the Earth's Radiant Energy System
CBH, CEH	Cloud base, effective height
CTH	Cloud top height
CC	CALIOP-CPR
CER, COD	Cloud particle effective radius, cloud optical depth
CET	Cloud effective temperature
CERES	Clouds and the Earth's Radiant Energy System
CF	Cloud fraction
CFI, CFW	Ice, water cloud fraction
CM	CERES-MODIS
CPR	CloudSat Cloud Physics Radar
C3M	CALIPSO, CloudSat, CERES, and MODIS
C6	MODIS Collection 6
Ed2, Ed4	CERES Edition 2, Edition 4
FAR, HR	False alarm, hit rate
H	Cloud geometrical thickness
HA	Horizontal averaging
HKSS	Hanssen-Kuiper's Skill Score
ID	Identification
IWP	Ice water path
JAJO	January, April, July, October
MAST	MODIS Atmosphere Science Team
MODIS	MODerate-resolution Imaging Spectroradiometer
NHM	Northern hemisphere midlatitude
NHP	Northern hemisphere polar
NOC	Non-opaque clouds
NP	Non-polar
OC	Opaque clouds
PO	Polar
POD	Probability of detection
PSC	Polar stratospheric cloud
RL-GEOPROF	Radar-Lidar Geometric Profile
SDD	Standard deviation of the differences
SIC	Snow/ice covered
SIF	Snow/ice free
SL, ML	Single-layer, multi-layer

This work was supported in part by the National Aeronautics and Atmospheric Administration through the CERES Project. *(Corresponding author: Christopher R. Yost)*

C. R. Yost, P. Minnis, S. Sun-Mack, Y. Chen are with Science Systems and Applications, Inc., Hampton, VA 23666 USA. (email: christopher.r.yost@nasa.gov)

W. L. Smith, Jr. is with the Science Directorate, NASA Langley Research Center, Hampton, VA 23681-0001 USA.

TOA	Top of atmosphere
B	Planck function
r	Correlation coefficient
T_{11}, T_s	11-μm brightness temperature, surface skin temperature
T_{low}, T_c	Low cloud, top temperature
ε	Cloud emissivity
$\Delta\tau$	Optical depth bias
τ, τ_{CA}	VIS cloud optical depth, CALIOP cloud optical depth
τ_{ci}	VIS cloud optical depth for detecting ice phase in ML clouds

I. INTRODUCTION

THE Clouds and the Earth's Radiant Energy System (CERES) Project [1] has been retrieving cloud properties from Terra (1030 LT equatorial crossing time) and Aqua (1330 LT equatorial crossing time) since March 2000 and July 2002, respectively. The primary measurements used by CERES for retrieving cloud properties are from the MODerate-resolution Imaging Spectroradiometer (MODIS; [2]) that operates on each satellite. These measurements and their conversion to physical parameters are part of a long-term climate dataset useful for many applications. The motivation for the products, and the algorithms used for the CERES Edition 4 (Ed4) cloud properties were described and discussed at length in Part I of this paper [3] along with some validation studies. Part II is devoted to more comprehensive evaluations of the accuracies of certain cloud parameters.

The Aqua platform is part of the Afternoon Constellation, better known as the A-Train, which consists of a group of satellites placed in orbit such that they continuously take measurements of the same locations on the Earth within a few minutes of each other. Data taken by active instruments on two of the A-Train satellites, CloudSat [4] and the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO, [5]), provide the most detailed cloud information available on a global basis, albeit limited to a narrow near-nadir curtain of the atmosphere. The Cloud Profiling Radar (CPR) on CloudSat penetrates through almost all clouds providing a profile of cloud particle reflectivity that can be used to accurately characterize cloud vertical structure. It can also be used to estimate cloud water content and water path [6]. Relatively insensitive to small cloud particles in low density clouds, the CPR profile often misses some cirrus clouds and the top portions of some convective systems. Those clouds are easily detected by the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) on CALIPSO. The CALIOP, which complements the CPR, is also used to detect aerosols and retrieve their microphysical properties as well as those of non-opaque ice clouds. These sensors have been used to assess imager-derived cloud properties since their launches in 2006. Those capabilities are exploited here to evaluate the retrievals of certain cloud properties retrieved from Aqua MODIS using the CERES Ed4 algorithms.

II. DATA

The basic assumption here is that the active-sensor data provide a ground truth for assessing the cloud properties retrieved or estimated from MODIS. All matched data from January, April, July, and October (JAJO) from 2015 and 2016 are used for phase comparisons, while JAJO 2010 data are used for all other comparisons. The different years used in the analyses arose as a matter of convenience. It is assumed that the results should be typical of any year.

A. CERES -MODIS (CM) Cloud Properties

This study uses Aqua CERES-MODIS (CM) single-pixel retrievals of cloud phase, cloud top height (CTH), cloud effective height (CEH), cloud base height (CBH), ice cloud thickness (H), and non-opaque ice cloud optical depth (COD) and ice water path (IWP). The nominal 1-km MODIS pixels are sampled from every other scan line and every fourth location along the scan line. Thus, the operational CM products do not provide a continuous field of cloud properties. Only those pixels classified as cloudy by the Ed4 cloud mask [7] or by the retrieval process itself [3] are considered here.

B. Validation Data

The CALIOP Release 4.20 Cloud Layers product [8,9] are used to assess CTH, thin cirrus optical depth [10] and IWP, cloud top thermodynamic phase [11], and cloud layering. The Cloud Layers product contains cloud information derived over a range of horizontal averaging resolutions: HA = 1/3 km, 1 km, 5 km, 20 km, and 80 km. The lower resolution products were developed by averaging the 1/3-km lidar backscatter intensity profiles over lengthening horizontal distances to detect very faint, very low COD clouds. The minimum detectable optical depth decreases with decreasing resolution. As in [7], the primary dataset used here is the 5-km Cloud Layers product and MODIS pixels were matched to each 5-km segment of the CALIPSO ground track. However, the 5-km Cloud Layers product lacks information for clouds detected at HA < 5 km, so the 1/3- and 1-km Cloud Layers products were also incorporated into the matching process so that clouds of all HA are represented in the resulting matched dataset, and subsequent analyses can be characterized by HA. Because the matched dataset includes the 1/3- and 1-km data, the cloud fraction, CF_{CA}, in a given 5-km CALIPSO footprint can range from 0 – 100%. The subscript CA indicates it is for CALIOP only. The cloud-layer top heights and bases detectable by CALIOP along with their retrieved optical depths are also included in the matched dataset. The cumulative CALIOP optical depth, τ_{CA}, represents the optical depth of all the layers having a retrieval. It is limited to values typically less than 3.0 because the lidar signal is usually attenuated at greater optical depths.

All MODIS pixels within a 2.5-km radius of the midpoint of each CALIPSO 5-km segment were considered to be spatial matches. Because of the CERES sampling of MODIS data, this results in only 1-4 MODIS pixels for a given CALIPSO 5-km pixel. In this study, any cloud layers detected only with the 20-km or 80-km averaging are, as a default, treated as being cloud free, unless otherwise noted. In general, COD is so small at those resolutions that the cloud is not detectable with passive imager data. To demonstrate the sensitivity of the retrievals to

the resolution of the CALIPSO product, however, some comparisons include the lower resolution data.

The CPR retrievals combined with the CALIOP data in the Radar-Lidar Geometrical Profile product (RL-GEOPROF, [12]) are employed to evaluate the cloud thickness H and base height CBH retrievals. The CPR can detect the bases of clouds with optical depths exceeding the transmission limit of the CALIOP. Hereafter, the RL-GEOPROF data are referred to as CALIOP-CPR, or CC, data. CC cloud base heights $CBH_{CC} < 1$ km are highly uncertain because of surface interference of the reflected radar signal. Displays of the cloud boundaries are generated from the most recently updated version of the CloudSat, CALIPSO, CERES, and MODIS (C3M) merged product [13]. C3M uses a special full-resolution CERES MODIS Ed4 analysis that does not use the 2 km x 4 km sampling of the operational product.

III. Comparisons

A. Cloud Phase

Uncertainties in cloud phase are examined by comparing the phase selections from Ed4 to those from CALIOP. As an initial assessment, mean liquid and ice fractions (global and zonal) were computed at a 3° resolution from the matched data for each CALIPSO HA resolution. When CALIOP indicated the presence of more than one phase in a column, the phase of the uppermost layer was used for cloud fraction computations. The geographic distributions of Ed4 water cloud fraction, *CFW* (Fig. 1a) show regional means that are very similar to those from CALIOP (Fig. 1b) when HA $\leq$ 5 km. Geographical distributions of the liquid (Fig. 1c) and ice phase (Fig. 1d) differences, however, suggest more disagreement than is readily apparent in the means. False liquid cloud tops are most frequent in the tropical convective zones, Mongolia and western China, and the northwestern Pacific. The best agreement occurs over the subtropical high-pressure domains. The patterns are similar for the ice cloud differences. During the day, the Ed4 retrievals significantly (> 0.05) underestimate ice cloud-top fraction over the continents where ice clouds are most likely to occur. Exceptions are found over Antarctica and central Greenland, where the Ed4 ice fraction is too high.

The phase of the uppermost layer varies with HA, so the CALIOP water and ice cloud fractions were computed as functions of HA. As demonstrated in the plots of zonal cloud fraction shown in Figs. 1e and 1f, CALIOP cloud fraction varies significantly depending on the HA scales considered. It is common to treat 80- and 20-km cloud detections as clear sky, and indeed Ed4 zonal means agree most closely to CALIOP when HA scales no larger than 5 km are considered (i.e., 20- and 80-km detections are treated as clear sky). Ed4 liquid water fraction is generally overestimated relative to CALIOP regardless of the HA scales considered while ice cloud fraction is underestimated (Figs. 1e and 1f). The greatest zonal biases in water cloud fraction (~0.06) occur between 30°N and 60°N. Water cloud fraction differences are small in magnitude south of 65°S and around 30°S and 20°N. Ed4 zonal ice cloud fractions (Fig. 1f) are smaller than CALIOP at all latitudes except south of 75°S over Antarctica. There is more variability in the CALIOP ice cloud fraction than water cloud fraction due to HA, particularly in the tropics where thin cirrus are abundant. The Ed4 overestimate (underestimate) of liquid (ice) clouds relative to those from CALIOP increase with increasing HA, mainly because more thin cirrus clouds are detected using longer HA scales.

At night (Fig. 2), the patterns in mean *CFW* from Ed4 (Fig. 2a) are quite similar to those from CALIOP (Fig. 2b), except for areas poleward of 60° latitude. Overestimates of liquid water cloud fraction (Fig. 2c) in the tropics are smaller than during the daytime (Fig. 1c), but polar (PO, poleward of 60° latitude) *CFW* is significantly underestimated. Mean Ed4 *CFW* is quite close to the 1 and 5-km HA CALIOP averages over most of the nonpolar (NP, equatorward of 60° latitude) zone (Fig. 2e), but is significantly underestimated in the PO zones. Conversely, Ed4 tends to overestimate ice cloud fraction *CFI* over the poles (Figs. 2d and 2f) and falls short of CALIOP's ice fraction in tropics. The global mean nocturnal Ed4 *CFI* is nearly unbiased (-0.006) relative to CALIOP data for HA $\leq$ 5 km. The daytime bias (-0.058) is 0.052 more negative during the night. Thus, more ice clouds are mistaken as liquid during the day. For liquid water clouds, the nighttime difference is -0.047 compared to +0.033 during the day. At night, the ice false alarm rates are nearly triple the daytime values.

More detailed comparisons of cloud phase were performed at the pixel level by organizing the phase outcomes into contingency tables as illustrated in Table I and various statistical metrics were computed to quantify the outcomes. This analysis only used data for which CALIOP and MODIS detected a single phase and both cloud fractions were 1.0. Phase mismatches are expected in scenarios involving multi-layer or multi-phase clouds. The overcast, single-phase restriction helps to eliminate phase mismatches in heterogeneous scenes where a single-phase choice is insufficient to accurately characterize the scene. The statistical metrics computed include the number of pixels N and the fraction of cloudy pixels used; the hit rate, $HR = (a + d) / N$, and false alarm rates for both water and ice clouds ($FAR_w = b / [b + d]$ and $FAR_i = c / [a + c]$, respectively) based on the definitions used by [14]; and the Hanssen-Kuipers' skill score (*HKSS*) defined in [15]. *HKSS* ranges in value from -1 to 1, with values ≤ 0.0 indicating no skill in making the correct determination. The phase bias, defined as $PB = (c - b) / N$, was also determined. A positive bias occurs when the number of incorrect ice outcomes exceeds the number of incorrect water outcomes ($c > b$).

Table II summarizes the JAJO 2015-2016 comparisons of the 100% cloudy, single-phase CALIOP pixels for both PO and NP snow/ice free (SIF) land and water surfaces as well as all snow-ice-covered (SIC) surfaces combined. The analyzed daytime and nocturnal cases represent about 68% and 70% of the matched 100% cloudy pixels, respectively, for those local time periods. Phase selection agreement (hit rate) is best over NP ice-free ocean during the day (0.971) and decreases to ~0.92 over snow-free land and over all SIC areas, producing an overall value of 0.951 for all surface types. During the day, the algorithm tends to slightly overestimate ice clouds over SIF ocean and over all SIC areas. The bias (positive for too much ice) is consistent with the higher ice false alarm rates. Over SIF land, the Ed4 phase algorithm favors liquid water phase

selections. For all scenes together, the bias of 0.004 is slightly tilted to ice, while the HR or phase agreement fraction is 0.95. For all of the daytime cases, regardless of surface type, HKSS $\geq$ 0.89. It reaches 0.95 for the nonpolar ocean case indicating very high skill in the Ed4 phase determination. Due to the fundamentally different spatial resolution and sampling methods of CALIOP (70-m beam) and MODIS (nominal 1-km pixels), it is very unlikely that the CALIOP and MODIS phase will always agree, and thus HR and HKSS are expected to be < 1.

Global HR drops to 0.871 at night as a result of 0.06 and 0.12 reductions in fraction correct over SIF and SIC areas, respectively. At night, ice phase is over-detected everywhere with an overall bias of 0.10 for ice clouds. The ice FAR is greatest at 0.34 for clouds over SIF PO ocean, but it is quite high over SIF PO land areas also. The water FAR is relatively small for all scenes except NP land. At night the classification skill drops compared to daytime, as might be expected from *HR*. The *HKSS* ranges from 0.52 over SIC regions to 0.85 over SIF nonpolar ocean.

The decrease in phase discrimination skill at night is likely due to the loss of information from the solar channels, reduced contrast between surface and cloud temperatures, and, especially over PO regions, the increased frequency of supercooled water clouds. The impact of supercooled clouds is potentially significant as demonstrated in Fig. 3, which shows the nighttime cloud phase probability distributions as a function of the cloud effective temperature CET for 2016 Aqua retrievals over the northern hemisphere polar (NHP, 60°N-90°N) and northern hemisphere midlatitude (NHM 30°N-60°N) zones. Fifty percent of the NHM water clouds have CET > 270 K, compared to only 15% for NHP water clouds, owing to the much colder environment in polar regions. The sharp peak in the NHP water distribution around 257 K is in contrast to the much-reduced frequency of liquid clouds at lower temperatures. The results in Table II and Fig. 2e suggest that a significant fraction of the NHP clouds identified as ice with CET between 233 and 260 K are actually supercooled liquid clouds. As indicated by [16] and references therein, many of those NHP liquid water clouds identified as ice by Ed4 are probably mixed phase clouds with water on top and ice underneath.

The results in Table II are representative of overcast single-layer, single-phase observations only and so represent only a subset of the entire matched dataset. Cloud phase validation metrics are affected by a number of factors including CALIOP HA, multi-layer or multi-phase cloud systems, and partially cloudy scenes. Fig. 4 summarizes the variation of *HR*, *HKSS*, and phase bias with these factors. Results are shown for two broad categories: "all overcast and partially cloudy data" (left half) and "overcast data only" (right half). These two categories are each subdivided into "all phase retrievals", i.e., including multi-layer (ML) and mixed-phase scenes (MP), and "single-layer, single-phase only". In the case of ML/MP systems, the phase of the uppermost cloud layer determines the cloud phase used in the analysis. For each category the phase metrics were plotted for different surface conditions as functions of CALIOP HA. The results in the rightmost column of Figure 4 are identical to those presented in Table II. Of the factors considered here, ML cloud systems have the most significant impact on the metrics. *HR* increases by as much as 0.15 (Figs. 4b,d) when ML clouds (Figs. 4a,c) are excluded from the analysis and *HKSS* increases by as much as 0.33 (Figs. 4e-h). This is not surprising since ML clouds introduce ambiguity to the cloud phase selection. As in Table II, Ed4 phase is generally biased towards the ice phase for "single-layer, single-phase" data (Figs. 4i-l). However, when all data are considered the bias shifts towards the water phase which is consistent with the cloud fraction values in Figs. 1 and 2. This result suggests that Ed4 retrieves the water phase in a significant portion of multi-layer systems in which thin cirrus is the actual top layer, even after discounting the 80- and 20- km detections.

The CALIOP HA scales considered in the analysis have a significant impact for most surface types but to a lesser extent compared to the impact of ML/MP clouds. HA has an even smaller impact when only overcast data are considered. For most surface conditions, the highest *HR* and *HKSS* are achieved for HA ≤ 1 and the values decrease when cloud detections using larger HA are included. Including 5-km detections results in only modest decreases in *HR* and *HKSS*, but including 20- and 80-km detections decreases *HR* and *HKSS* by up to ~0.04 and ~0.07, respectively. Partially cloudy data have similar effects on the metrics, reducing *HR* and *HKSS* by ~0.03 and ~0.07, respectively. This is probably a result of partially cloudy MODIS pixels being interpreted as high cloud. *HR* and *HKSS* are generally higher for daytime (solid lines) than nighttime (dashed lines). The daytime phase bias is towards water clouds in most cases, particularly when "all phase retrievals" are considered, and this result is consistent with Fig. 1. The nighttime phase is generally biased towards ice clouds especially for polar and snow/ice-covered surfaces, consistent with results shown in Fig. 2.

B. Cloud vertical structure

Fig. 5 shows the C3M active sensor cloud boundaries (orange) compared with CBH_{Ed4} (blue dots) and CTH_{Ed4} (black dots) derived from Aqua-MODIS data taken 2 October 2009. Fig. 5a-5e show daytime retrievals while Fig. 5f shows nighttime data. The Ed4 cloud top height includes the thick ice cloud correction that was mistakenly overwritten in the operational Ed4 code [3] and should be applied to ice clouds having COD > 8. In Fig. 5a, a nearly 5-km thick cirrus cloud overlies a marine stratus deck over the tropical ocean. The Ed4 thickness is ~ 5 km, but the top is too low, likely because the cirrus is optically thin and the underlying stratus optical depth is larger, affecting CET and hence the CEH estimate. As the cirrus thins, gradually exposing the stratus, the Ed4 top and base drop until CTH_{Ed4} matches up with the CC top. Further in time, CTH_{Ed4} edges slightly above the CC top before matching it again near the end of the segment. The CBH_{Ed4} is initially too low, but agrees with the CC base after minute 18. A few scattered false high clouds, seen around minute 19, are likely in the vicinity of a cloud edge, which the algorithm interprets as ice clouds.

The vertical boundaries of moderately vigorous tropical convection and scattered precipitating cumulus clouds are plotted in Fig. 5b. Here, the cumulus CTH_{Ed4} generally agree with CC, but the edge pixels are exceptions. For the larger

precipitating cumuli (minutes 31-32) CBH_{Ed4} is typically higher than its CC counterpart. In these cases, it is difficult to distinguish between the actual cloud base and precipitation below cloud base. Similarly for the deeper convective cloud, CBH_{Ed4} is typically higher by 1 km or so, and rain obscures the actual cloud base. CTH_{Ed4} is typically within 1 km of CC, but sometimes is overestimated by 2 km. The extremely deep convection in Fig. 5c is topped with a thin cirrus veil detected only by CALIOP. CTH_{Ed4} follows the denser cloud top, underestimating the actual top by ~1 km in the core and more where the veil is separate from the denser clouds below. Precipitation obscures the actual cloud base, so it is not clear how accurate CBH_{Ed4} is for many of the clouds in this case.

A mid-latitude example of both marine stratus and relatively deep convection is shown in Fig. 5d. In this case, both the low-level and thick clouds are outlined fairly well, while some of the cirrus cloud boundaries are either under- or overestimated. A 10-km thick system over the southern polar ocean (Fig. 5e) is poorly characterized with the top too low and base too high. CTH_{Ed4} for the stratus around minute 34.5 is severely underestimated for a low cloud. Conversely, at night (Fig. 5f) the cloud boundaries of a 6-km system are reasonably matched up until minute 47 when the overlapped cloud effect occurs. Near minute 48, the cirrus CTH_{Ed4} is too low, probably due to a very low COD. A section of false clouds is detected near the Antarctic continent, while some very thin cirrus clouds are detected over its plateau. The top heights are underestimated by 1-2 km and the base heights are severely underestimated. The false clouds arise from the poor thermal contrast and low predicted surface temperature accuracy that occur over many polar areas [7].

These examples highlight the complex nature of the scenes and some of the errors that are encountered in reconstructing cloud vertical structure from passive sensor data. The estimated cloud top and base height parameters all depend on the cloud effective height, which relies on the retrieved cloud effective temperature. In turn, CET is affected by the accuracies of the surface temperature, lapse rate, and cloud particle optical properties. To better quantify the errors and determine the reliability of the cloud height parameters, several comparisons are conducted using several different data groupings.

1) Cloud top height

Cloud top heights are compared using only the CALIOP 5-km resolution data that are overcast and for which the topmost clouds are either entirely liquid or ice, unless otherwise indicated. The matching MODIS data must also be overcast and one phase. The results are divided into three categories: snow/ice-free ocean and land, and snow/ice-covered. The last category includes both ocean and land surfaces. Since the difference between CEH and CTH_{Ed4} is minimal, only cloud-top height is compared for liquid clouds. For ice clouds, both CEH and CTH_{Ed4} are compared with CALIOP since the effective and top height differences can be significant. The comparisons are performed separately for opaque and non-opaque clouds, where opaque refers to the absence of a return lidar signal from the surface. It typically occurs for clouds having COD > 2 or so. Single-layer (SL) clouds are first examined and then the overall uncertainties are determined relative to the upper-most cloud top height.

a) Liquid cloud tops: Fig. 6 compares the daytime JAJO 2010 Aqua Ed4 and CALIOP SL, single-phase liquid cloud top heights. The mean heights for CALIOP are given by <ZCAL> and the standard deviation of the differences as SDD. Over SIF ocean (Fig. 6a), the Ed4 and CALIOP heights are highly correlated (r = 0.87) and the differences are 0.00±0.67 km. A greater portion of the water clouds over SIF land are found above 3 km (Fig. 6b). In these cases, the correlation is the same as that over water surfaces, but CTH_{CM} averages 0.10±0.91 km less than that from CALIOP. The increased bias over SIC areas (Fig. 6c), 0.21±1.00 km, is accompanied by a drop in correlation to r = 0.75. Combining results over all surfaces yields a global daytime mean difference of 0.03 ± 0.77 km for SL water clouds (Fig. 7a). During the night (Fig. 7b), Ed4 overestimates SL water CTH by 0.14±0.71 km (Fig. 7b). The increased bias at night is mostly a result of overestimates over ocean, where the bias is 0.15±0.63 km (not shown). The differences over land and SIC are smaller at night than during the day.

CTH comparisons were also performed on three different categories of liquid-phase clouds of progessively increasing scene complexity. The first category (A) includes only SL liquid clouds as determined by CALIOP. The second category (B) includes SL and ML cloud systems, but all layers are liquid phase, e.g, mid-level liquid-phase cloud overlying shallow cumulus. The last category (C) includes SL and ML systems with potentially mixed-phase (MP) or dual-phase (DP) conditions, but in all cases the top cloud layer is liquid phase. In other words, each category inlcudes SL liquid clouds (e.g., Fig. 7a) and progressively more complex scenes were added to comprise categories B and C. The expectation is that CTH differences will be smallest for category A and will increase for categories B and C due to ML and DP complications. Results for category C are reprentative of a large dataset, ("all" liquid clouds) while results for category A represent a relatively small dataset. The results for all liquid clouds over all surfaces combined are plotted in Figs. 7c and 7d for day and night, respectively. Roughly, 577 x 10^3 points are added to the results in Fig. 7a to obtain the plot in Fig. 7c for all of the daytime matched liquid clouds. The additional matches are mostly underestimates of CTH_{CA}, indicated by an expansion of the points below the identity line (or one-to-one line). The result is that for all liquid clouds, the mean difference is -0.32 ± 1.24 km, a significant change in bias and standard deviation. For nighttime data, the inlclusion of ML/DP matches increases the number of points seen in Fig. 7b by ~386 x 10^3, but the impact is similar to that during the daytime. The bias and SDD drop to -0.12 km and 1.07 km, respectively.

Table III summarizes the comparison statistics for SL and all liquid clouds for the different surface types. For all scenes during the daytime, the bias decreases and the SDD increases when ML and multi-phase conditions are added to the SL-only cases. The magnitude of the bias, however increases as its value decreases for the SIF land and ocean surfaces. The multiphase, liquid-top clouds account for ~15% of the total number of overcast liquid water cloud conditions and so are a relatively small portion of the dataset. At night, the bias and SDD have a similar dependence on the type of cloud conditions used in the

comparison, but the fraction of multiphase liquid top clouds drops to 10%. It is clear that CTH_{Ed4} is most accurate for homogeneous liquid clouds.

b) Ice cloud tops: For comparisons with Ed4 ice clouds, the CALIOP data were used to separate the clouds into opaque and non-opaque, where the opaque cloud (OC) appears to be a single-layer, contiguous ice cloud with no backscatter return from the surface or a lower layer cloud. The CALIOP beam is completely scattered before or at the level of the ice cloud base. This limit corresponds to an optical depth of 0.5 – 5.5 during the day and 3 – 9 at night [10] depending on the actual scattering properties of a given ice cloud. Non-opaque clouds (NOC) are defined as those having a return signal from the surface or an underlying water cloud.

Fig. 8 shows the scatterplots of SL opaque ice CTH_{CA}, and three cloud heights from the matched Ed4 data: left: CEH; middle: CTH_{Ed4} as archived; and right: CTH_{CM} computed from CEH using the formula (equation 6 of [3]) that should have been applied to the archived data. These results include all surface types. On average, CEH is 2.3 km less than CTH_{CA} during the day (Fig. 8a) and 2.0 km below CTH_{CA} at night (Fig. 8d). The Ed2 correction, mistakenly applied to produce the archived Ed4 product, yields $CTH_{Ed4} - CTH_{CA}$ differences of -1.7 and -1.4 km during the day (Fig. 8b) and night (Fig. 8e), respectively. Applying the Ed4 correction to CEH *ex post facto* drops the $CTH_{CM} - CTH_{CA}$ differences by a factor of 2 or better (Figs. 8c and 8f). SDDs are ~1.3 km in the daytime and ~1.7 km at night. The correlation is higher during the day ($r \sim 0.86$) than at night ($r \sim 0.82$). The horizontal features seen around 12 and 15 km correspond to model tropopause heights that limit the retrieved and adjusted heights. These artifacts are more evident in the CTH plots because the height adjustments are limited to be equal to or less than the tropopause height in most cases, one notable exception being deep convective overshooting cloud tops.

The opaque ice cloud difference statistics are summarized in Table IV for categories similar to to those used to compare liquid clouds. Note that Category C includes all opaque columns even if the uppermost cloud layer was not opaque, and that a thin cirrus cloud overlying an opaque water cloud may be characterized as an optically thick ice-phase cloud by Ed4 and thus qualify for the cloud-top height parameterization for optically thick ice clouds. The daytime results in Fig. 8a respresent 57% of the 746 x 10^3 matched ice cloud in category (C). As for the liquid clouds, the magnitudes of the biases and SDDs for (B) increase relative to those for (A), and likewise for all OCs (C) relative to (B). On average, the corrected CTH_{CM} over land is closest to CTH_{CA} during the day, while at night, the corrected ocean mean is nearest to CTH_{CA}.

The SL non-opaque ice cloud effective and top heights are compared for day and night in Figs. 9 and 10, respectively. The number of samples is greatly reduced compared to the OC category, as are the correlation coefficients. Over land, the daytime biases and SDDs are smallest for both CEH (Fig. 9a) and CTH (Fig. 9d), while the ocean (Figs. 9b, 9e) and SIC (Figs. 9c, 9f) biases and SDDs are comparable to each other. The parameterization used in Ed4 to estimate CTH from CEH for optically thin ice clouds (e.g., NOC) is based on equations (24) and (25) of [17]. For the data in Fig. 9, that parameterization determines that, on average, CTH is ~1.3 km higher than CEH. But, placing the cloud top height 1.3 km above the apparent radiating height is insufficient to align CTH_{CM} with CTH_{CA}, leaving biases of -1.2 km over land and less than -2.1 km over other surfaces. Except for the clusters of points for $CTH_{CA} > 13$ km over SIF surfaces, the two datasets appear to be well correlated. Over SIC surfaces the same high-altitude clusters are not evident, but large differences exist for $CTH_{CA} < 11$ km.

The SL ice NOC correlations and biases are much improved at night over land and ocean (Fig. 10). The thin cloud CTH adjustment has been applied to most of the data, however, some Ed4 clouds qualified for correction with the thick cloud correction, which was applied to them retroactively and is included in all of the NOC CTH statistics. Over ocean, the CEH bias is only -1.36 km (Fig. 10b) compared to -3.4 km during the day, while *r* has doubled. This improvement leads to a CTH_{Ed4} underestimate of only 0.4 km (Fig. 10e). The SDDs also drop substantially. At -0.9 and -0.2 km, the respective nocturnal CEH (Fig. 10a) and CTH (Fig. 10d) biases over land are also substantially diminished relative to daytime, although the SDDs are similar to those during the day. Over SIC regions (Figs. 10c and f), the nighttime SDDs exceed their daytime counterparts although the biases, while slightly improved relative to daytime, are still quite large and *r* is little changed from its low daytime value. At night, the CALIOP heights over SIC areas include polar stratospheric clouds (PSCs) that are evident in Figs. 10c and 10f at altitudes exceeding 13 km. PSCs form during polar winter when the stratospheric temperatures fall below -78°C [18] and thus do not appear in Fig. 9. Evidently, some PSCs or their underlying clouds are detected by the Ed4 cloud mask, but the retrieved CET is either too high or, if it is colder than the tropopause temperature, is reset to the tropopause temperature. Thus, any CTH_{Ed4} corresponding to a PSC is rarely above 13 km.

Table V summarizes these results and those for the other NOC categories which include ML and multi-phase conditions as in Table IV. When the SIF and SIC results are combined, the global mean NOC SL CTH is underestimated by -1.95 km during the day and -0.97 km at night. In this averaging scenario, the PO regions, which are sampled much more frequently than other areas, are weighted heavily. Using areal averaging, assuming that the SIC areas correspond to ~14% of the surface, the global nocturnal mean SL CTH difference would be -0.53 km, while the daytime difference would essentially be unchanged. As the restrictions on matching are loosened, the daytime bias magnitudes and SDDs actually decrease slightly relative to their SL counterparts during daytime, while they generally increase at night. Overall, the NOC biases and SDDs exceed their OC counterparts in magnitude.

So far, the height comparisons have focused on overcast CM and CA pixels having the same phase for various categories. To get an overall assessment of the cloud top height differences regardless of phase, layering or coverage, the differences in CTH for all matched data were computed and averaged for three different CALIOP height categories: all, low ($CTH_{CA} < 5$ km), and high ($CTH_{CA} > 5$ km). Fig. 11 plots the resulting probability distributions of ($CTH_{CM} - CTH_{CA}$) for

JAJO 2010 for overcast (CF = 1.0) and mostly cloudy (CF ≥ 0.5) scenes. For the overcast cases (Fig. 11a), the low cloud difference peaks at 0 km with the bias magnitude (0.06 km) and SDD (1.07 km) being less than the loosest liquid phase matching category in Table III. The high cloud mode is at -0.5 km, while the distribution itself is highly skewed to negative values resulting in a large negative mean and SDD. Together with the high clouds, the mean daylight difference is -1.36 ± 2.74 km. The low cloud bias increases and becomes positive for mostly cloudy cases (Fig. 11b) and SDD increases as well. The high cloud differences are little changed because the majority of partially cloudy scenes are low-level water clouds. The histogram characteristics are negligibly changed despite a 24% increase in the number of points. Overall, the mean difference is -1.15 ± 2.88 km.

At night, the absolute mean differences decrease overall for both the overcast (Fig. 11c) and mostly cloudy (Fig. 11d) cases, but the SDDs rise relative to their daytime counterparts. The cores of the high cloud histograms broaden, while the fraction of extremes diminishes. At night, the inclusion of mostly cloudy points in Fig. 11d has a larger impact on the low-cloud differences than during the daytime. The difference in the means is twice that seen during the daytime. The positive low cloud biases suggest that some of the low clouds identified by CALIOP were mistaken as ice clouds by Ed4 since those pixels would not have been included in Table II. Reduced biases for the nocturnal high clouds relative to the daytime values reflect the differences seen for the NOC clouds. The lower nighttime biases for all clouds together result from the increased low cloud differences.

If only NP regions are considered for day and night together, the overall mean differences fall between the global day and night values. For overcast cases (Fig. 11e), the low cloud bias is 0.00 ± 0.98 km, which is smaller in mean and standard deviation than the corresponding global values, day or night. The high cloud bias and SDD lay between the global day and night values. For the mostly cloudy, NP cases (Fig. 11f), all of the biases and SDDs are between the global day and night values. The low-cloud SDD is nearly twice that of the overcast cases. The impact of the polar regions on the global mean differences appears to be greatest on overcast low cloud systems. This is not readily apparent in Table II, which does not include the height differences for clouds having incorrect phase assignments. Additionally, in the higher latitudes, NOC ice clouds can occur more often below 5 km than over nonpolar areas, e.g. Figs. 9f and 10f. Since their top heights are more uncertain than liquid clouds, the overall uncertainty increases with their inclusion in the low cloud statistics.

The above presentation focuses on the quantitative assessment of the various cloud top height errors, but does not directly show how the average cloud top vertical profiles from Ed4 compares to that from CALIOP. Fig. 12 plots the normalized frequencies of cloud height occurrence for CEH, CTH_{CM}, and CTH_{CA} for the JAJO 2010 period for all times of day. In the tropics (Fig. 12a), CEH and CTH_{CM} track CTH_{CA} quite well for clouds below 3 km, but are too frequent between 3 and 6 km, where influence of the ice cloud differential between CTH and CEH becomes apparent. Overestimation of CTH frequency diminishes to zero at 12 km before jumping to 0.02 again at 13.5 km. Underestimates occur between 13.75 and 15.25 km and above 16 km. The two CTH_{Ed4} peaks above 12 km, evident in the earlier scatterplots, seem to exaggerate two small but perceptible CALIOP local maxima near the same altitudes. Over the midlatitudes (Fig. 12b), the profiles, though compressed due to the shallower troposphere, are similar up to ~9 km, where the high cloud maximum begins. Above that level, both CEH and CTH_{CM} underestimate the frequency of high clouds. Ed4 low and midlevel cloud occurrence is overestimated in the PO regions (Fig. 12c), while the high cloud (> 9 km) frequency is too low.

2) Cloud thickness

Single-layer cloud thicknesses determined from a combination of CC data are compared with their matched Ed4 counterparts in Fig. 13 for JAJO 2010. During daytime, the mean water cloud thickness difference, $H_{Ed4} - H_{CC}$, is -0.25 ± 0.83 km (Fig. 13a). Although many of the points cluster around the identity line, an arm of points extends horizontally centered around H_{Ed4} ~ 1.2 km with values of H_{CC} reaching nearly 10 km. The correlation coefficient is 0.44. At night (Fig. 13b), the correlation decreases to 0.25 along with the bias to -0.42 ± 0.73 km. The distribution of points at night is similar to those seen during the day, except that the peak value of H_{Ed4} is only 1.5 km instead 2.2 km. This sharp maximum at night is due to the default limits placed on water clouds at night. The daylight maximum is also sharp due to the limit of COD = 150 used in Ed4 for all clouds.

The Ed4 ice cloud thickness parameterization was tested using only one month of CC data (see Fig. 2 of [3]) before being implemented. To examine how well it works operationally, the thicknesses for opaque and non-opaque ice clouds are compared separately using the 4 months of 2010 data. During daytime (Fig. 13c), the opaque cloud thickness is biased by -0.24 ± 2.14 km, with the differences distributed fairly symmetrically about the identity line. This bias is the same as that for water clouds but it is only 3% of mean cloud thickness, instead of 25%. The correlation, $r = 0.79$, is significantly greater than the water cloud value. At night (Fig. 13d), the H_{Ed4} values flatten out at around 5 km with a secondary plateau at ~7.5 km, revealing significant thickness underestimates. Larger H_{Ed4} values up to 14 km also occur but they correspond to H_{CC} between 14 and 18 km. The correlation remains significantly greater than that seen for daytime water clouds. Overall, the mean nocturnal difference is -2.74 ± 2.51 km at night. The greater errors at night are a result of the limitations of retrieving COD and effective particle radius CER using only infrared radiances.

For non-opaque ice clouds, the points are reasonably well distributed around the identity line during the day (Fig. 13e), except for H_{CC} between 0.5 and 3.0 km, where the parameterization produces underestimates. This results in a lower correlation coefficient, 0.47, and a worse bias, -0.37 ± 1.51 km, than for opaque clouds. Similar underestimates at the low end are seen at night (Fig. 13f), but less frequently than during the day. Some of the flattening around 5 km seen in Fig. 13d also occurs at night with additional underestimates. Values of H_{CC} extend to 18 km for these optically thin clouds,

presumably corresponding to the PSCs discussed earlier. Overall, the mean nocturnal difference is -1.04 ± 2.11 km and r = 0.38.

3) Cloud base height

The Ed4 cloud base height is the computed as the difference between CTH_{Ed4} and H_{Ed4}. Fig. 14 shows the scatterplots for JAJO 2010 liquid cloud base heights over all surfaces for SL clouds (left) and for all cases in which the uppermost layer was a liquid cloud (right). The number of samples is reduced in each case compared to those in Table II because the RL-GEOPROF was unable to obtain a reliable cloud base in all instances. All four of the scatterplots show similar features, such as a moderately-dense cluster of points along the identity line and a more dense cluster of points between CBH_{CC} = 0.5 and 1.5 km that extends up to CBH_{Ed4} = 7 km with diminishing number density. Those points mirror the horizontal patterns seen in the water cloud thickness comparisons. In all 4 cases, CBH is overestimated. During the day, the SDDs exceed their CTH counterparts, but are slightly smaller at night. The bias magnitudes are smallest for daytime SL clouds (Fig. 14a) and greatest for all nocturnal clouds (Fig. 14d). Adding the ML/MP water clouds to the SL case during the day (Fig. 14b) slightly increases both the bias and SDD. A similar effect occurs at night when ML/MP clouds are added to the SL clouds (Fig. 14c). For SL clouds, the overestimates of CBH are expected given the negative biases seen in Figs. 13a and 13b and the slight overestimates of CTH. For the all-cloud cases, the CBH_{Ed4} error is larger, as would be expected when there are gaps between the cloud layers.

Scatterplots of cloud base heights from CERES and RL-GEOPROF are presented in Fig. 15 for all opaque SL/ML/MP ice clouds (category C) in Table III. CBH_{Ed4} was computed based on the archived CTH. Again, because of the inability to find bases for all of the clouds, the numbers of points are less than those in Table III. The correlation (r = 0.83) during the day (Fig. 15a) is nearly as high as that for CTH. The linear cluster of points is skewed below the identity line but offset by a feature at CBH_{CC} ~1 km that is similar to the one in Fig. 14. This feature is enhanced at night (Fig. 15b), when the correlation decreases and Ed4 is unable to obtain many thick cloud estimates because of the optical depth constraints of using only infrared for the retrievals. On average, Ed4 underestimates CBH by 0.5 km during the day and overestimates by 1.8 km at night relative to CC. The SDDs are larger than their CTH counterparts in Table III. The daytime underestimate is due to several factors, particularly the negative bias in the archived CTH_{Ed4}. If CTH_{Ed4} were corrected in Fig. 15a, CBH_{Ed4} would rise by an average of 0.70 km and the mean bias in CBH_{Ed4} would be reduced to -0.2 km, consistent with the thickness bias in Fig. 13c.

The non-opaque ice cloud bases from the two cloud products in Fig. 16 show significant correlation for both day and night datasets. For SL NOCs, r increases from 0.74 in the daytime (Fig. 16a) to 0.80 at night (Fig. 16c). The cluster of points falling well below the identity line for CBH_{CC} > 10 km in Fig. 16a corresponds to a similar feature over SIF surfaces in Figs. 9a and 9b, but is absent in Fig. 16c. At night, the magnitude of the bias drops by 1.3 km from its daytime value. For all NOC cases, the absolute mean difference decreases by 0.8 km from day (Fig. 16b) into night (Fig. 16d), while the correlation remains at ~0.81. Underestimates during the day are not surprising given the large negative biases in CTH_{Ed4}, which are not offset by the daytime bias in NOC thickness. At night, the positive biases are mainly due to two factors: the COD limitations and the larger thickness bias, which is comparable to the CTH bias in Table IV. During the night, the bias for SL ice is -0.13 km over SIC areas (not shown), compared to 0.63 km over SIF surfaces. This contrasts with the daytime results, which show a CBH bias of -2.35 km over SIC surfaces, compared to -1.4 km for SIF areas. Overall, the magnitudes of the CBH biases and SDDs (~2.3 km) are smaller than their CTH counterparts in all four categories in Fig. 15.

C. Ice cloud optical depth and water path

Optical depths retrieved for single-layer NOC clouds from JAJO 2010 Aqua MODIS and CALIOP are compared in Fig. 17 for day (top) and night (bottom). The CALIOP optical depths include those from both constrained and unconstrained retrievals. Over SIF land during the day (Fig. 17a), the points are clustered in a quasi-linear fashion with a relatively steep slope above the identity line, yielding a 96% overestimate compared to COD_{CA}. At night (Fig. 17d), the correlation more than doubles and the bias drops to 15%. Over ocean, the daytime bias (Fig. 17b) is 46% and the correlation is better than the nocturnal correlation over land. The bias drops to 5% at night over ocean (Fig. 17e), while r nearly doubles from its daytime value. Over SIC areas, the ice cloud COD is severely overestimated during both day (Fig. 17c) and night (Fig. 17f). The minimal correlation between the CC and Ed4 values is negative, indicating that the Ed4 NOC retrievals over SIC surfaces have no value and should not be trusted. Similar results were found when using only the constrained CALIOP retrievals.

Fig. 18 compares NOC ice water paths for SIF areas only. The mean IWP_{Ed4} is roughly half its CALIOP counterparts during both day and night. Over ocean, the correlation coefficient rises from 0.30 during the day (Fig. 18a) to 0.71 at night (Fig. 18c) at night, while the bias changes from -54% to -44%. The land correlation coefficient during the day (Fig. 18b) more than doubles at night (Fig. 18d), even though the corresponding underestimates of IWP rise from 48% to 56%. These underestimates are surprising given the overestimates of COD in Fig. 17.

IV. Discussion

The cloud parameters examined above are all interrelated in some fashion as cloud phase is critical for selecting the model for retrieving the effective temperature and optical depth, parameters used to estimate cloud height. Here, the results are further examined to better understand the differences between the Ed4 Aqua and A-Train active sensor retrievals, to determine how the Ed4 relative accuracy compares to other retrieval methods, and to open avenues for improving future retrieval algorithms.

A. Cloud Phase

When the cloud-top phase is clearly defined, as in Table I for both Ed4 and its matched CALIOP counterpart, the Ed4

phase retrievals are quite reliable over SIF areas, being correct 90% and 96% of the time at night and during the day, respectively. Including the SIC areas yields the corresponding global average hit rates of 87 and 95%. The latter value can be compared to the 92% hit rate found for the MODIS Atmosphere Science Team (MAST) Collection 6 single-phase matches with Version-3 CALIOP 1 and 5-km data taken during November 2012 [19]. Ed4 phase hit rates similar to those in Table I and Fig. 4 were obtained using CALIOP Version 3 data [20], so it is unlikely that there are any significant cloud-top phase differences between the Version 3 and 4 CALIOP data. The differences between the MAST and Ed4 phase comparisons with CALIOP are likely due to algorithmic and/or seasonal sampling differences.

Unfortunately, for ease of remote sensing, clouds are not confined to a single-phase within a given atmospheric column. Figs. 2, 3, and 4 clearly demonstrate that single-phase comparisons do not represent the whole picture. When the phase conditions are more ambiguous for a given pair of matched retrievals, the accuracy of the Ed4 phase selection retrieval accuracy diminishes. To determine the conditions when the phase selections disagree, the CALIOP data were divided into 13 scene categories and the dominant Ed4 cloud scene ID (clear, water, ice, no-retrieval) was assigned to the scene category of its matching CALIOP pixel. The fractions of each Ed4 scene were then computed for each CALIOP category.

Fig. 19 summarizes the JAJO 2015-16 results for ocean scenes (top) and all scenes together (bottom). For clear ocean CALIOP pixels, the Ed4 scene ID is 95% clear and 5% cloudy during the daytime (Fig. 19a) and 91% clear at night (Fig. 19b) with false cloudy pixels split between water and ice. For the CALIOP water cloud types, the best daylight agreement over ocean is 96% for SL water clouds, while the worst is 88% for SL water-dominant clouds. At night, the multi-phase water-top clouds produce the lowest congruence at 67%; SL clouds remain the best identified at 88%, although the ML water clouds are close, at 87%. The SL ice clouds are classified properly by Ed4 at the 95-96% level at any time of day, while the greatest misclassifications are 58% and 35% for non-opaque cirrus over water clouds during day and night, respectively. Being in the third most populated cloud category, these Ed4 phase disagreements for non-opaque cirrus over water clouds account for 61% and 48% of all day and night phase differences, respectively.

For all surface types, the clear statistics in Figs. 19c and 19d for day and night, respectively are nearly the same as the ocean results. The SL water cloud accuracy during the day is at 94%, but at night drops to 79%, presumably from the inclusion of the polar night cases. Errors in classifying SL water-dominant and multi-phase water top clouds as liquid are a little worse than their ocean counterparts for all times of day. Identification of SL ice clouds as ice phase is diminished when the land and SIC data are added to the histograms with 9% and 14% of the pixels misclassified for day and night, respectively. Similar to the ocean cases, the cloud type presenting the greatest challenge for identifying the top as ice is the non-opaque cirrus over liquid water clouds. During the day, this category accounts for 60% of the cloud phase misclassifications, while at night it contributes 29% of the discrepancies. This drop in the contribution of multilayer clouds to the nighttime phase differences relative to the ocean results is due not just to a decrease in errors for multilayer clouds, since the in-cloud-type percentages differ by only 6%. Rather, it is mostly due to a rise in errors for other cloud types, especially SL water, an impact felt mostly over SIC regions at night.

To further examine the largest error source in determining the cloud top phase, the frequencies of Ed4 ice and water phase were determined as a function of the semi-transparent cirrus cloud optical depth when such clouds occurred over liquid water cloud layers for JAJO 2015-2016. Fig. 20 plots the Ed4 phase frequency (water: blue, ice: red) as a function of the upper-layer ice cloud optical depth τ_{CA} determined from CALIOP profiles. The overlaid blue bar chart gives the normalized frequencies of τ_{CA} for 0.1 intervals of optical depth, while the green line indicates the fraction of the CALIOP ML profiles classified as clear by Ed4. The top and bottom plots use HA $\leq$ 5 km and HA $\leq$ 80 km, respectively. Of all the scenes identified as "non-opaque ice over water" ML systems in Fig. 19, ~5% had τ_{CA} deemed to be of poor quality (i.e., extinction quality control flag > 1) and were therefore excluded from this analysis, so these results represent ~95% of the cases. In these cases, the water phase dominates for low τ_{CA}, but the frequency of Ed4 ice-phase retrievals increases with increasing τ_{CA} and ice becomes the dominant phase when τ_{CA} exceeds some threshold τ_{ci}. During the day (Fig. 20a) for HA $\leq$ 5 km, when $\tau_{CA} < 0.1$, Ed4 classifies 5% of the scenes as clear, 19% as ice, and 76% as water. The misclassification decreases such that there is an equal likelihood of classifying the scene as an ice or water cloud when $\tau_{CA} \sim 0.9$. As COD increases, the likelihood of calling the scene water decreases to ~20% for $\tau_{CA} = 3.0$, but the number of profiles also diminishes rapidly. The classification statistics are similar over SIF water and SIC surfaces, with τ_{ci} occurring around 0.8, while over land $\tau_{ci} \sim 1.2$ (not shown).

At night for HA $\leq$ 5 km (Fig. 20b), Ed4 is more likely to classify these ML scenes as ice. For $\tau_{CA} < 0.1$, Ed4 identifies the clouds as ice 29% of the time, while the false clear percentage is twice the daytime value. In this case $\tau_{ci} = 0.3$, indicating much less sensitivity to the underlying water cloud than during daytime. In that same vein, the percentage classified as ice for a given τ_{CA} interval rapidly increases to 90% at $\tau_{CA} = 0.8$, while during the day, it never exceeds 80%. The classification dependency in Fig. 20b is typical of the SIF surfaces. Over SIC areas, more false clear areas are diagnosed, but $\tau_{ci} = 0.15$, possibly due to the typically colder nature of the lower level clouds (not shown).

If the low-resolution averaging is included, the relative fraction of upper-layer ice clouds having $\tau_{CA} < 0.1$ is substantially increased, from 0.05 to 0.28 during the day (Fig. 20c) and from 0.11 to 0.36 at night (Fig. 20d). Incorporating the ML systems having the most tenuous cirrus layers by including the HA $\leq$ 80 km cases, decreases the fraction of ice classifications and raises the false clear detections for the first several τ_{CA} intervals, but has no impact on the crossover points. It increases the number of ML samples by 32% and 26% during day and night, respectively. The higher fraction of clear cases

found with the greater averaging length occurs for all surface types. This suggests that not only is the upper cloud very thin, but in many cases the lower-level cloud is also quite tenuous given the hit rates for cloud detection over ocean as a function of total CALIOP optical depth [7].

These differences in the phase selection between day and night are primarily driven by the relative contribution of the upper and lower clouds to the visible reflectance and infrared signals. During the day, the COD τ is based on the reflectance which, in the ML classification used here, is likely dominated by the lower-level cloud, while the 11 and 3.7-μm brightness temperatures affect the phase selection. In a simplified form, the retrieved cloud effective temperature is

$$CET = B^{-1}\{[B(T_{11}) - (1-\varepsilon)B(T_s)] / \varepsilon\}, \quad (1)$$

where T_{11} is the observed brightness temperature, T_s is surface temperature, and B is the Planck function at 11 μm and B^{-1} is its inverse. For near-nadir views, the cloud emissivity can be approximated as

$$\varepsilon = 1 - e^{-\tau/2}. \quad (2)$$

CER and *CET* are the main determinants of the phase selection and *CET* must be 273 K or less for the cloud phase to even be considered as ice. Thus, for example, if τ = 5 from the reflectance retrieval for a hypothetical ML system having, T_s = 300 K, *CET* = 220 K, and τ_{CA} = 0.2, and a low cloud temperature of T_{low} = 290 K, then T_{11} ~ 284 K. Using COD in (2), the emissivity would be ~0.92 and the retrieved value of T_c would be ~282.5 K from (1). In this instance, there would be no chance for the cloud to be classified as ice. As the upper cloud COD increases, the chances for ice selection increase, but there must also be an ice model solution for the observed 3.7-μm radiance. In most cases, T_c must be < 250 K for ice to be selected (e.g., Fig. 1).

At night, an iterative technique using three infrared channels is employed in Ed4 to simultaneously retrieve *CET*, COD, and particle size. That method is more likely to return an ice phase selection because it relies on the cloud emissivities at the different wavelengths. The difference between the assumed, T_s = 300 K, and the actual, 290 K, background temperatures is relatively small compared to the difference between T_s and upper-layer cloud temperature. Thus, the iteration will overestimate COD for the upper layer cloud, but its value will be considerably smaller than COD = 5 retrieved during the daytime. This can be quantified from the nocturnal results in Table V. Without the ML cases (category B) over ocean, the bias for NOCs is -0.37 km. Adding in the ML cases (category C) doubles the number of samples and results in a bias of -0.88 km. Thus, the mean ML height bias over ocean at night is -1.41 km, an error that is smaller than the SL ice cloud bias during daytime over ocean. Improving the phase selection for ML clouds during daytime would require using an infrared-only technique, similar to that used during the day, or to employ an explicit ML cloud detection and retrieval system.

B. Cloud vertical structure

1) Cloud-top height

The differences in Table II for liquid water clouds are similar to the 2 months of comparisons reported by [21]. During daytime, the Table II SL land and global SIC results are slightly worse than their counterparts in [21], but the SL differences are a little better for ocean scenes. At night, the current results have a bias of 0.15 km over ocean compared to 0.10 km from [21], with nearly the same SDD. However, the bias in category (B) in Table II is only 0.05 km. The SL nocturnal bias over land is near zero in [21], but 0.09 km here with similar SDDs. The differences between the current and those from [21] are likely due to the differences in sampled time periods, a different definition of overcast (only 2/3 of CALIPSO 333-m pixels had to be cloudy in [21]), and the use of 1-km matches in [21] instead of the 5 km used here. Overall, the conclusions reached by [21] are confirmed here in the differences and the vertical profile plots.

For the most part, the regional lapse rate approach to converting effective temperature to cloud-top height has been successful in reducing uncertainties in low cloud CTH to values below those found using a number of other techniques [21]. Yet, additional improvement should still be sought. The scatterplots in Fig. 7 indicate that matched SL water cloud heights are not evenly distributed around the identity line. The clusters of points found between CTH_{CA} ~ 1.5 km and CTH_{Ed4} ~2.5 km suggest some systematic flaw in the approach, which should be identified and addressed in future editions. For all water clouds (Figs. 7c,d), the ML clouds discussed above cause the broad areas of CTH underestimation. These can be addressed most effectively by explicitly identifying and interpreting ML cloud systems in the retrieval algorithms.

Ice cloud effective and top heights are too low for both opaque ice clouds, even though the *ex post facto* correction brings CTH_{CM} more in line with its CALIOP counterparts than the old Ed2 corrections used inadvertently for the Ed4 processing. If the intended Ed4 correction were applied to all of the opaque clouds (τ > 3 or so) in Fig. 8 instead of just for τ > 6, CTH_{CM} would be still be biased low by ~0.4 km, on average, during the day, but would be nearly unbiased at night. During the daytime, it is likely that more ML clouds are included in the ice cloud population than at night because the visible channel reflectance will yield a much higher COD for a ML system than the ice cloud by itself. At night, the COD retrieval is affected less by the presence of underlying low clouds. Again, better identification of ML clouds would improve the classification and height assignment even for "opaque" ice clouds. Other artifacts of the algorithm occur mainly at night, for example, the assignment of CTH_{Ed4} to particular levels and the severe underestimates of PSC heights. The former will require investigation of tropopause height constraints. Although not expected to be classified with the opaque clouds, PSCs are evident in the lower half of Fig. 8. As low-COD clouds, those in this category must have occurred over some thicker tropospheric clouds that raised the CALIOP column backscatter to saturation.

For NOCs, the Ed4 roughened ice-crystal model [22] yields lower optical depths for NOCs than the Ed2 smooth

crystal model, but the change is still insufficient to increase CEH enough to obtain a CTH_{Ed4} match with CALIOP during the day (Fig. 9). Several factors affect the estimate of CTH_{Ed4}: the surface temperature, COD, and the relationship between CEH and CTH. Surface temperature is known best for ocean areas, yet the differences between CTH_{Ed4} and CTH_{CA} are largest during the day. At night, the differences are significantly decreased because COD_{Ed4} is, on average, very close to CALIOP value. The formula used to relate CEH and CTH is the same day or night, so it cannot be a major error source. The retrieved COD is significantly overestimated during the day (Fig. 17) and relatively small biases in COD can produce large errors in CEH and hence CTH. To illustrate, Fig. 21 plots the underestimate in CEH computed as a function of COD for three different percentage overestimates $\Delta\tau$ of COD for a hypothetical cloud having CET = 220 K and surface temperature of 290 K. For a given $\Delta\tau$, the CEH underestimate increases with decreasing COD, while for given COD the underestimate increases non-linearly with rising $\Delta\tau$. The mean daytime CTH bias over ocean for NOCs -2.11 km for a mean COD of 0.77. If the conditions stipulated in Fig. 21 held for the results in Fig. 9b, a value of $\Delta\tau \sim 19\%$ would be expected instead of the 43% seen in Fig. 17b. The results in Fig. 21 are only for demonstrative purposes and represent only one of the many pairs of CEH and T_s that occur in reality. The actual variation of the height bias with $\Delta\tau$ changes with the differences between CET and T_s. Nevertheless, it is clear that the retrieved COD is the main source for the daytime CTH bias.

Although it is relatively successful at night, the parameterization to estimate CTH from CEH [23] could use additional improvement as the points at heights above 13 km at the bottom of Figs. 9 and 10 are mainly below the identity line. The parameterization is based on a relatively small number of measurements of cirrus clouds in the midlatitudes [24] for temperatures above 215 K and may not be representative of the lower temperatures seen at high altitudes in the tropics and in the midlatitudes during summer. Lower temperatures will, in general, produce diminished ice water contents and require a greater thickness for a given optical depth.

The Ed4 – CALIOP CTH differences in Fig. 11e or 11f can be compared to similar statistics made using the MAST C6 Aqua retrievals (MYD08) and CALIOP for August 2006. Frey et al. [25] found that for NP regions, the mean difference between MYD08 and CALIOP is -2.69 km with SDD = 3.86 km. For mean CALIOP cloud tops at altitudes $\leq$ 5 km, the difference is -0.55 km ± 1.05, while for clouds above 5 km, it is -4.41 ± 4.41 km. The absolute biases and SDDs are considerably less than those in Fig. 11e and mostly less than those in Fig. 11f. Minnis et al. [3] found that the mean nonpolar MYD08 cloud top heights for October 2015 are 1.21 km less than those from Ed4. That difference is consistent to within 13% of the average discrepancy of 1.40 km between the MAST-CALIOP and Ed4-CALIOP differences. The greatest discrepancy, 2.18 km for high clouds, seems surprising given the improved ice model employed by the MAST [26]. Since the MAST CTH relies on the CO_2-slicing method and infrared channels during all times of day [27], the retrieved visible COD does not impact the height retrieval as it does for the Ed4 algorithm.

The CERES system is designed foremost to relate clouds to the TOA radiation budget and secondarily to the surface radiation budget. Thus, it is important that the retrieved cloud properties are as radiatively consistent as possible with the CERES radiative fluxes. Such consistency is realized when zero differences are found between the CERES TOA and surface flux retrievals with radiative transfer calculations using input from the CM cloud retrievals and other auxiliary information [28]. To achieve that consistency, accurate COD and cloud radiating heights are especially important. Reaching that goal is relatively straightforward when the clouds in question are SL systems, but in the absence of explicit ML cloud information, the problem is more complicated. Interpreting a ML cloud scene with the single-layer assumption used here, should yield an essentially correct radiating height because the 11-μm radiance proportionally combines the emitted radiation from the water and ice cloud layers. Additionally, it probably returns a relatively accurate cloud optical depth for a majority of systems because the influence of the ice cloud on the retrieved value is small. Thus, achieving the correct phase and height of the overlying ice cloud is not likely to have a great impact on radiative flux estimates at the TOA and, possibly, at the surface. If the correct ice cloud phase and top height for a ML system were retrieved and still assumed to be from a SL cloud, i.e., the retrieved column COD corresponds to the high cloud, then the TOA longwave fluxes would be significantly underestimated in most cases. This can be understood in light of the example given in the previous subsection.

While the current SL-algorithm approach is reasonably satisfactory for TOA and surface radiation computations, CERES has the additional objective of computing within-atmosphere heating rates. To attain a more accurate realization of that goal, it is necessary to have information about cloud layering and thickness [29]. A ML algorithm was implemented in the Ed4 processing, but it proved to be unreliable. Thus, more accurate and dependable ML detection and retrieval algorithms are still needed to accompany the CERES flux measurements.

2) Cloud thickness and base height

The SL liquid cloud thickness and base height comparisons in Figs. 13 and 14 show significant numbers of points clustered about the identity line demonstrating good correlation for part of the dataset. The correlation is diminished in each set of comparisons which have a branch of data extending vertically or horizontally from the correlated clusters. All of the observations for the liquid-cloud thickness parameterization used in both Edition 2 and Ed4 were made in marine stratus and stratocumulus conditions at a single location in the subtropical northeastern Pacific Ocean, where the clouds had mean effective temperatures of ~13°C [30]. Since it is based on a single variable, the retrieved column optical depth, it implicitly assumes that liquid water content (LWC) is constant. LWC typically decreases with temperature [31], thereby requiring a greater thickness at low temperatures for a given value of COD than for higher temperatures. Since no supercooled clouds were included in the thickness parameterization development, it is likely that the thickness in colder liquid clouds is underestimated.

To explore this possibility, the differences between the Ed4

and CC liquid cloud base heights were compiled as a function of latitude for JAJO 2010. To minimize the impact of inhomogeneities in the cloud fields, the analysis used only those pixels that were elements of liquid-only cloud systems at least 25 km in horizontal extent. Fig. 22 shows the zonal mean differences and SDDs separately for clouds having CET > 273 K (red) and those with CET $\leq$ 273 K (blue). The near-zero means for the warm clouds are accompanied by small SDDs. Over land (Fig. 22a) and ocean (Fig. 22b) the respective mean differences are -0.14 ± 1.04 km and 0.04 ± 0.47 km. When only supercooled clouds are considered, the differences and SDDs are considerably larger, particularly over the tropics.The land and ocean mean differences are 0.49 ± 1.04 km and 0.39 ± 0.81, respectively. The SL water cloud-top heights have similar mean differences to those for the warm clouds indicating that the thickness parameterization works very well for warm clouds, but not for supercooled clouds. From the observations in [31], it appears that the mean LWC at 13°C is representative of warm clouds, which is consistent with the limited conditions used for the parameterization. To obtain more accurate thickness estimates for both warm and supercooled liquid clouds, a temperature-dependent parameterization will need to be developed.

It is clear from Fig. 13c that the Ed4 ice cloud thickness parameterization is a significant advance over its Ed2 counterpart, but it can use some additional improvement, especially for thin ice clouds as demonstrated in Figs. 13e,f. It is not particulary effective at night for opaque clouds, mainly because of the inability to consistently retrieve large optical depths without solar channels. New methods for estimating thick ice cloud optical depths at night [32] may help alleviate that shortcoming. As noted earlier, underestimation of opaque cloud base height is mostly attributable to the negative bias in the archived Ed4 cloud-top heights. Some of the overestimates for CBH_{CC} < 2 km in Fig. 15a, however, may be an artifact of the CloudSat profile, which does not discriminate between cloud and precipitation. Thus, the actual cloud base may be higher than indicated by the RL-GEOPROF profile for opaque clouds.

3) Cloud optical depth and water path

To improve the daytime thin cloud optical depths, it is necessary to either use an infrared-only method such as that employed for nocturnal retrievals or use a more representative ice crystal model. For radiative consistency, however, the infrared and visible reflectance retrieval models should produce the same COD values. Thus, an improved ice crystal model is warranted. One example of a more representative ice crystal model is one based on aggregates of severely roughened ice crystals that was adopted by the MAST for their C6 retrievals [26]. It has asymmetry factors of ~0.75 for most particle effective radii that result in better radiative consistency with CALIOP and infrared retrievals [33]. The CERES Ed4 model has asymmetry factors between 0.77 and 0.80, hence, it yields greater COD values than would arise from the MAST model. Replacement of the roughened ice crystal column model with one having a smaller asymmetry factors, such as the two-habit model introduced by [34] and tested by [35], should be a priority for future CERES editions.

The Ed4 ice water paths were found to substantially underestimate the reference CALIOP values. This result was unexpected given that an earlier comparison using July 2013 CALIOP Version 3 data from [20], and shown here in Fig. 23, yielded relatively good agreement between the Ed4 and CALIOP IWPs. For that Version 3 comparison, the IWP_{Ed4} means are 25% and 17% larger during the day than their IWP_{CA} counterparts over ocean (Fig. 23a) and land (Fig. 23c), respectively. At night, the ocean (Fig. 23b) and land (Fig. 23d) Ed4 means are 42% and 0% greater than the respective CC averages. The daytime results are more in line with the COD overestimates in Fig. 17, while the nighttime bias over ocean is surprising. Since IWP_{CA} is the product of COD, CER, and a constant, the value of CER determines the IWP for a given COD. Platnick [36] showed that at nadir, a retrieval using 3.7-μm radiances should yield a value of CER representative of the entire cloud for COD < 2.2. For these clouds, then, CER_{Ed4} should be a reasonable representation of the cirrus clouds analyzed in Figs. 17 and 22, particularly when it was retrieved directly and did not use the default value of 10 μm [3]. The median, non-default daytime value of CER_{Ed4} is ~32 μm, while the mean for all NP areas is 27 μm. Since COD is overestimated during the day, the retrieved value of CER_{Ed4} would have to increase by more than a factor of 2 for IWP_{Ed4} to match IWP_{CA} in Fig. 17. This is confirmed by [37], who noted that the Version 4 CER_{CA} average is 76.5 μm. Ed4 retrieves values exceeding 70 μm less than 2% of the time. The main difference between CALIOP IWP in Versions 3 and 4 is the use of a new temperature-dependent parameterization of ice water content and particle size [38] for Version 4 [39] instead of the parameterization of [40] for the older version. While the relatively good agreement with the Version 3 results may simply be fortuitous, the differences between the Ed4 and CALIOP Version 4 retrieved particle sizes are troublesome. If the latter are the correct values, then retrievals of ice-cloud CER using 3.7-μm or 2.1-μm radiances must be severely biased as most current algorithms have ice CER median values between 30 and 40 μm [41]. The differences between the CALIOP Version 4 IWP estimates and those from passive retrievals for semitransparent ice clouds must be reconciled to ensure that measurements of CER for ice clouds using both passive and active sensors are trustworthy.

V. Summary and Conclusions

Several key CERES Ed4 cloud properties have been compared to state-of-the-art cloud parameter retrievals from the CALIOP and CPR active instruments. Relative to similar previously published comparisons, the analysis is comprehensive. It used a variety of spatial resolutions and matching criteria to understand the sensitivities of imager retrievals to ice cloud optical depth and to better identify the main sources of uncertainties in each of the parameters. Overcast CALIOP pixels from the 5-km horizontal averaging product were assumed to comprise the standard reference, since other studies have used similar matching assumptions and clouds detected at larger averaging lengths are difficult to detect in passive imagery. Comparisons using other horizontal averaging lengths and fractional cloud cover were also

performed to include nearly all of the Aqua MODIS matches with the CALIOP data.

For single-phase CA pixels, the Ed4 and CALIOP cloud top phase agreed 95.8% and 92.0% of the time over snow-free and snow/ice-covered areas, respectively, during the day. At night, the corresponding hit rates decreased to 89.9% and 79.8% due to loss of solar reflectance information. During the day, liquid water cloud fraction was slightly overestimated over snow-free regions, mainly due to retrievals over land. Too many ice clouds were estimated over snow/ice areas. At night, ice clouds were identified too frequently over all areas, especially over polar regions. Cloud phase agreement generally increased using the 1-km CALIOP data, but decreased when horizontal averaging exceeded 5 km. Disagreement between the phase selections increases when all clouds are considered and the dominant phase of the scene determines the CALIOP phase.

The primary source of phase selection discrepancies is ice-over-water multilayer cloud systems. During the day, the Ed4 algorithm selects liquid cloud for 50% or more of these cases when the ice cloud optical depth is < 0.9. This 50% threshold drops to 0.3 at night, probably due to the nighttime algorithm's tendency to favor the ice-phase retrievals as discussed previously. Adding the multilayer clouds detectable by CALIOP only at 20 and 80-km averaging increases the number of misclassifications, but does not change the thresholds because the ice cloud optical depths of the added pixels are well below the threshold value. Only 2-6% of overcast water clouds are mistaken as ice clouds during the day, but at night, the fraction increases to 9-16% resulting in an overall ice-phase bias at night for the matched pixels. Thus, phase selection could be improved if pixels with multilayer clouds can be explicitly identified and the interpretation of infrared channel radiances could be refined to discriminate between supercooled liquid and ice clouds at night. While more improvement is desirable, the Ed4 phase selection for single phase clouds is accurate at the 90+% level compared to CALIOP in all but the worst conditions of polar night, where phase discrimination and cloud detection are difficult for all algorithms (e.g., [3], [19]).

Compared to CALIOP, Ed4 cloud-top heights for single-layer water clouds differ by 0.03 ± 0.77 km and 0.14 ± 0.70 km during the day and night, respectively. Including the remaining 35% and 28% water-dominant cloud types for day and night, respectively, the corresponding differences are -0.32 ± 1.24 km and -0.12 ± 1.07 km. The rise in uncertainty, when all water dominant clouds are considered, is due to the inclusion of multilayer water clouds and mixed-phase clouds and pixels. Ice-cloud top heights are underestimated, on average.

For single-layer opaque ice clouds, the global differences during day and night are -0.74 ± 1.29 km and -0.71 ± 1.64 km, respectively, if the Ed4 thick ice cloud correction is applied to the effective heights. Without the *ex post facto* correction, the corresponding archived heights are 1.75 and 1.37 km too low compared to CALIOP. The correction was intended to reduce the bias to zero, but optical depth restrictions for applying the correction and multilayered clouds introduce additional error when the correction is applied to operational data. Explicit determination of multilayered clouds systems and expanding the application of the top-height parameterization to clouds with COD > 3 would bring the Ed4-CALIOP top height differences much closer to zero both day and night.

Semi-transparent, single-layer ice cloud heights are relatively accurate (-0.30 ± 1.90 km) at night over snow-free areas, when only infrared channels are used in the retrieval. The agreement for the same types of scenes is worse, at -1.88 ± 2.53 km, during the day. This difference is primarily attributable to the reflectance model used to determine the cloud optical depth, which, in turn, affects the cloud effective height retrieval. More accurate clear-sky reflectances and the use of an ice-cloud reflectance model based on cloud particles with lower asymmetry factors for all sizes would reduce the error in the non-opaque cloud-top heights.

Cloud thickness is estimated from empirical formulae and then subtracted from the cloud top height to estimate cloud base height. Liquid cloud base heights were found to be as accurate as the cloud -top heights when the cloud effective temperature exceeds the freezing point, but are too high for colder clouds. This finding is consistent with the drop in liquid water content in clouds as the temperature decreases and it indicates that the cloud thickness parameterization should be temperature dependent. Both opaque and non-opaque ice cloud thicknesses are more accurate than cloud-top height for the same clouds during the day, with differences of -0.24 ± 2.14 km and -0.37 ± 1.51 km, respectively. The errors increase dramatically at night, particularly for opaque ice clouds, owing to the infrared blackbody limits on optical depth retrievals. As a result, the cloud base height errors generally mirror the cloud top height errors during the day and are worse at night. It may be possible to improve the ice cloud thickness parameterization further by focusing on the extrema, the thinnest and thickest clouds, which the parameterization tends to over- and under-estimate, respectively.

Compared to the CALIOP retrievals, non-opaque ice-cloud optical depths are overestimated by ~60% during the day and by 8.7% at night. Very little skill is evident in the retrievals over snow-covered areas. The larger daytime error is due to the ice-cloud reflectance model discussed above. Cloud ice water path for non-opaque ice clouds is underestimated by about 50% compared to the CALIOP Version 4 mean, but is overestimated by ~25% compared to the CALIOP Version 3 average. The latter is consistent with the overestimation of the ice cloud optical depths. If the Version 4 mean is correct, the Ed4 ice particle effective radius would need to be a factor of two larger than its current value, yet it is within ±20% of all other passive retrievals. Further evaluation of this discrepancy is needed to determine if the parameterization employed by CALIPSO in its latest version is suitable. If it is, then all retrievals of ice cloud effective radius using infrared, shortwave infrared, and near-infrared radiation must be revaluated.

A variety of new methods have been developed over the past 5 years to address some of the outstanding sources of uncertainty and directly reduce the errors in estimating some of the parameters discussed here. In particular, artificial neural network approaches have shown some promise for improved estimates of cirrus cloud-top height and optical depth [42] and estimates of opaque ice cloud optical depth at night [32]. The neural network technique of [43] appears to provide the most accurate cloud-top height to date, while multi-layer cloud

detection and upper-layer ice cloud top and base height estimates have been greatly improved using a different neural network method [44]. These and other advances must be considered for future editions of the CERES cloud retrieval system. It should be noted that while cloud-top height is critical for knowing cloud vertical structure, cloud effective height is most important for computing TOA radiation. Thus, any new approach must consider both the top and effective heights as they are distinctly different for ice clouds and possibly some liquid water clouds. It is possible, using certain assumptions about the vertical distribution of the cloud water content and an accurate estimate of the cloud phase and optical depth to reasonably estimate the cloud effective radiating height. Addition of effective particle size and cloud base height retrievals should enhance the accuracy of any such estimate.

This study has quantified the uncertainties in several cloud properties retrieved from Aqua MODIS data by the CERES Ed4 retrieval system. All results are based on measurements taken near the nadir view and may not necessarily be valid for measurements taken at higher viewing zenith angles. The sensitivity of using different CALIOP averaging lengths and pixel cloud fractions highlight the need to carefully select and describe the A-Train data used in any comparisons with imager cloud property retrievals, so that assessments of relative accuracy among the various retrievals are performed in the same manner. Overall, the findings here and in [3] indicate that the Ed4 cloud properties, at least over snow-free areas, are comparable to and sometimes more accurate than those from other global passive cloud retrieval algorithms. The uncertainties provided by the comparisons in this study should be valuable when using the Ed4 cloud data for weather and climate research.

ACKNOWLEDGMENT

The authors would like to thank M. Vaughan, A. Garnier, and other CALIPSO team members for discussions of the CALIPSO version 4 product. The CERES, CALIPSO, and C3M data are available at the NASA LaRC Atmospheric Sciences Data Center (**https://eosweb.larc.nasa.gov**). The RL-GEOPROF data were obtained from CloudSat Data Processing Center at the Cooperative Institute for Research in the Atmosphere (http://www.cloudsat.cira.colostate.edu)**.**

REFERENCES

[1] B. A. Wielicki, B. R. Barkstrom, E. F. Harrison, R. B. Lee III, G. L. Smith, and J. E. Cooper, "Clouds and the Earth's Radiant Energy System (CERES): An Earth Observing System Experiment," *Bull. Amer. Meteor. Soc.*, vol. 77, 853-868, 1996.

[2] W. L. Barnes, T. S. Pagano, and V. V. Salomonson, "Prelaunch characteristics of the Moderate Resolution Imaging Spectroradiometer (MODIS) on EOS-AM1," *IEEE Trans. Geosci. Remote Sens*., vol. 36, pp. 1088-1100, 1998.

[3] P. Minnis, S. Sun-Mack, C. R. Yost, Y. Chen, W. L. Smith, Jr., F.-L. Chang, P. W. Heck, R. F. Arduini, Q. Z. Trepte, K. Ayers, K. Bedka, S. Bedka, R. R. Brown, E. Heckert, G. Hong, Z. Jin, R. Palikonda, R. Smith, B. Scarino, D. A. Spangenberg, P. Yang, Y. Xie, and Y. Yi, "CERES MODIS cloud product retrievals for Edition 4, Part I: Algorithm changes to CERES MODIS," *IEEE Trans. Geosci. Remote Sens.*, 2019, submitted.

[4] G. L. Stephens, et al., "CloudSat mission: Performance and early science after the first year of operation," *J. Geophys. Res.*, vol. 113, D00A18, 2008, doi:10.1029/2008JD009982.

[5] D. M. Winker, M. A. Vaughan, A. Omar, Y. Hu, K. A. Powell, Z. Liu, W. H. Hunt, and S. A. Young, "Overview of the CALIPSO mission and CALIOP data processing algorithms," J. *Atmos. Oceanic Tech.*, *26*, pp. 2310-2323, 2009. doi:10.1175/2009JTECHA1281.1.

[6] R. T. Austin, A. J. Heymsfield, and G. L. Stephens, "Retrieval of ice cloud microphysical properties using the CloudSat millimeter-wave radar and temperature," *J. Geophys. Res.*, vol. 114, D00A23, 2009, doi:10.1029/2008JD010049.

[7] Q. Z. Trepte, P. Minnis, S. Sun-Mack, C. R. Yost, Y. Chen, Z. Jin, F.-L. Chang, W. L. Smith, Jr., K. M. Bedka, and T. L. Chee, "Global cloud detection for CERES Edition 4 using Terra and Aqua MODIS data, *IEEE Trans. Geosci. Remote Sens.*, vol. 57, 2019, doi:10.1109/ TGRS.2019.2926620.

[8] M. A. Vaughan, K. A. Powell, R. E. Kuehn, S. A. Young, D. M. Winker, C. A. Hostetler, W. H. Hunt, Z. Liu, M. J. McGill, and B. J. Getzewitch, "Fully automated detection of cloud and aerosol layers in the CALIPSO lidar measurements," *J. Atmos. Ocean. Tech.*, vol. 26, pp. 2034–2050, 2009, doi.10.1175/2009JTECHA1228.1.

[9] M. A. Vaughan, M. Pitts, C. Trepte, D. Winker, P. Detweiler, A. Garnier, B. Getzewich, W. Hunt, J. Lambeth, K.-P. Lee, P. Lucker, T. Murray, S. Rodier, T. Tremas, A. Bazureau, and J. Pelone, "Cloud-Aerosol LIDAR Infrared Pathfinder Satellite Observations (CALIPSO) data management system data products catalog, Release 4.10," NASA Langley Research Center Document PC-SCI-503, Hampton, Va., USA, 2016.

[10] S. A. Young, M. A. Vaughan, J. L. Tackett, A, Garnier, J. B. Lambeth and K. A. Powell, 2018: "Extinction and optical depth retrievals for CALIPSO's Version 4 data release", *Atmos. Meas. Tech.*, vol. **11**, 5701-5727, 2018. doi:10.5194/amt-11-5701-2018.

[11] Y. Hu, D. Winker, M. Vaughan, B. Lin, A. Omar, C. Trepte, D. Flittner, P. Yang, B. Baum, W. Sun, Z. Liu, Z. Wang, S. Young, K. Stamnes, J. Huang, R. Kuehn, and R. Holz, "CALIPSO/CALIOP cloud phase discrimination algorithm," *J. Atmos. Oceanic Tech.*., vol. 26, pp. 2293-2309, 2009.

[12] G. G. Mace and Q. Zhang, "The CloudSat radar-lidar geometrical profile (RL-GeoProf): Updates, improvements and selected results," *J. Geophys. Res.*, vol. 119, pp. 9441-9462, 2014, doi:10.1002/2013JD021374.

[13] S. Kato, F. G. Rose, S. Sun-Mack, W. F. Miller, Y. Chen, D. A. Rutan, G. L. Stephens, N. G. Loeb, P. Minnis, B. A. Wielicki, D. M. Winker, T. P. Charlock, P. W. Stackhouse, K.-M. Xu, and W. Collins, "Improvements of top-of-atmosphere and surface irradiances with CALIPSO, CloudSat, and MODIS-derived cloud and aerosol properties," *J. Geophys. Res*., vol. 116, D19209, 2011, doi:10.1029/2011JD016050.

[14] K. G. Karlsson and E. Johansson, "On the optimal method for evaluation cloud products from passive satellite imagery using CALIPSO-CALIOP data: example investigating the CM SAF CLARA-A1 dataset," *Atmos. Meas. Tech.*, vol. 6, pp. 1271-1286, 2013, doi:10.5194/amt-6-1271-2013.

[15] F. Woodcock, "The evaluation of yes/no forecasts for scientific and administrative purposes," *Mon. Wea. Rev.*, vol. 104, pp. 1209-1214, 1976.

[16] H. Morrison, G. de Boer, G. Feingold, J. Harrington, N. D. Shupe, and K. Sulia, "Resilience of persistent Arctic mixed-phase clouds," *Nat. Geosci.*, vol. 5, pp. 11-17, 2011, doi:10.1038/NGEO1332.

[17] P. Minnis, S. Sun-Mack, D. F. Young, P. W. Heck, D. P. Garber, Y. Chen, D. A. Spangenberg, R. F. Arduini, Q. Z. Trepte, W. L.Smith, Jr., J. K. Ayers, S. C. Gibson, W. F. Miller, V. Chakrapani, Y. Takano, K.-N. Liou, and Y. Xie, "CERES Edition 2 cloud property retrievals using TRMM VIRS and Terra and Aqua MODIS data: Part I: Algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, 4374-4400.

[18] M. C. Pitts, L. R. Poole, and R. Gonzalez, "Polar stratospheric cloud climatology based on CALIPSO spaceborne lidar measurements from 2006-2017," *Atmos. Chem. Phys.*, vol. 18, pp. 10881-10913, 2018, doi://10.5194/acp-18-10881-2018.

[19] B. Marchant, S. Platnick, K. Meyer, G. T. Arnold, and J. Riedi, "MODIS Collection 6 shortwave-derived cloud phase classification algorithm and comparisons with CALIOP," *Atmos. Meas. Tech.*, vol. 9, pp. 1587-1599, 2016, doi: 10.5194/amt9-1587-2016.

[20] CERES (2016), CERES_SSF_Terra-Aqu_Edition4A Data Quality Summary, *CERES Project Document*, 30 June, 19 pp. https://eosweb.larc.nasa.gov/project/ceres/quality_summaries/CER_SSF_Terra-Aqua_Edition4A.pdf

[21] S. Sun-Mack, P. Minnis, Y. Chen, S. Kato, Y. Yi, S. Gibson, P. W. Heck, and D. Winker, "Regional apparent boundary layer lapse rates determined from CALIPSO and MODIS data for cloud height determination," *J. Appl. Meteorol. Climatol.,* vol. 53, pp. 990-1011, 2014,

doi:10.1175/JAMC-D-13-081.1.

[22] P. Yang, G. W. Kattawar, G. Hong, P. Minnis, and Y. X. Hu, "Uncertainties associated with the surface texture of ice particles in satellite-based retrieval of cirrus clouds: Part II. Effect of particle surface roughness on retrieved cloud optical thickness and effective particle size," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, pp. 1948-1957, 2008, doi:10.1109/TGRS.2008.916472.

[23] Minnis, P., P. W. Heck, and E. F. Harrison, 1990: The 27-28 October 1986 FIRE IFO Case Study: Cloud parameter fields derived from satellite data. *Mon. Wea. Rev.*, **118**, 2426- 2446.

[24] Minnis, P., D. F. Young, K. Sassen, J. M. Alvarez, and C. J. Grund, 1990: The 27-28 October 1986 FIRE IFO Case Study: Cirrus parameter relationships derived from satellite and lidar data. *Mon. Wea. Rev.*, **118**, 2402 - 2425.

[25] R. Frey, B. Baum, A. Heidinger, S. Ackerman, B. Maddux, and P. Menzel, "MODIS CTP (MOD06) Webinar #7: Cloud top pressure, effective emissivity, and phase," *C6 Atmosphere Team Webinar Series*, 90 pp., 20 August, 2014. Available at https://modis-images.gsfc.nasa.gov/Webinar2014/MODIS_C6_Cloud_Top_Products_Menzel.pdf

[26] S. Platnick, K. G. Meyer, M. D. King, G. Wind, N. Amarasinghe, B. Marchant, G. T. Arnold, Z. Zhang, P. A. Hubanks, R. E. Holz, P. Yang, W. L. Ridgway, and J. Riedi, "The MODIS cloud optical and microphysical products: Collection 6 updates and examples from Terra and Aqua," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, pp. 502-525, 2017, doi:10.1109/TGRS.2016.2610522.

[27] B. A. Baum, W. P. Menzel, R. A. Frey, D. C. Tobin, R. E. Holz, and S. A. Ackerman, "MODIS cloud-top property refinements for Collection 6," *J. Appl. Meteor. Climatol.*, vol. 51, pp. 1145-1163, 2012, doi:10.1175/JAMC-D-11-0203.1.

[28] S. Kato, F. G. Rose, D. A. Rutan, T. J. Thorsen, N. G. Loeb, D. R. Doelling, X. Huang, W. L. Smith, W. Su, and S. H. Ham, "Surface irradiances of Edition 4.0 Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) data product," *J. Climate,* vol. 31, pp.4501-4527, 2018, doi:10.1175/JCLI-D-17-0523.1

[29] S. Kato, F. G. Rose, S.-H. Ham, D. A. Rutan, A. Radkevich, T. Caldwell, S. Sun-Mack, W. F. Miller, and Y. Chen, "Radiative heating rates computed with clouds derived from satellite-based passive and active sensors and their effects on generation of available potential energy," *J. Geophys. Res.,* vol. 124, pp. 1720-1740, 2019, doi:10.1029/2018JD028878.

[30] P. Minnis, P. W. Heck, D. F. Young, C. W. Fairall, and J. B. Snider, "Stratocumulus cloud properties derived from simultaneous satellite and island-based instrumentation during FIRE," *J. Appl. Meteorol.*, vol. 31, pp. 317-339, 1992.

[31] Gultepe, I. and Isaac, G. A., "Liquid water content and temperature relationship from aircraft observations and its applicability to GCMs," *J. Climate*, vol. 10, pp. 446-452,1998.

[32] Minnis, P., G. Hong, S. Sun-Mack, W. L. Smith, Jr., Y. Chen, and S. Miller, 2016: Estimation of nocturnal opaque ice cloud optical depth from MODIS multispectral infrared radiances using a neural network method. *J. Geophys. Res.*, **121**, doi:10.1002/2015JD024456.

[33] R. E. Holz, S. Platnick, K. Meyer, M. Vaughan, A. Heidinger, P. Yang, G. Wind, S. Dutcher. S. Ackerman, N. Amarsinghe, F. Nagle, and C. Wang, "Resolving ice cloud optical thickness biases between CALIPSO and MODIS using infrared retrievals," *Atmos. Chem. Phys.*, vol. 16, pp. 5075-5090, 2016, doi:10.5194/acp-16-5075-2016.

[34] C. Liu, P. Yang, P. Minnis, N. Loeb, S. Kato, A. Heymsfield, and C. Schmitt, "A two-habit model for the microphysical and optical properties of ice clouds," *Atmos. Chem. Phys.*, vol. 14, pp. 13719-13737, 2014, doi:10.5194/acp-14-13719-2014.

[35] N. Loeb, P. Yang, F. G. Rose, G. Hong, S. Sun-Mack, P. Minnis, S. Kato, S.-Ham, and W. L. Smith, Jr., "Impact of ice cloud microphysics on satellite retrievals and broadband flux radiative transfer calculations," *J. Climate*, vol. 31, 2018, pp. 1851-1864, doi:10.1175/JCLI-D-7-0426.1.

[36] S. Platnick, "Vertical photon transport in cloud remote sensing problems," *J. Geophys. Res.*, vol. 105, pp. 22919-22935, 2000, doi:10.1002/2000JD900333.

[37] M. Vaughan, NASA Langley Research Center, Hampton, VA, private communication, July 2019.

[38] A. J. Heymsfield, C. Schmitt, and A. Bansemer, "Ice cloud particle size distributions and pressure-dependent terminal velocities from in situ observations at temperatures from 0° to -86°C," *J. Atmos. Sci.*, vol. 70, pp 4123-4154, 2013, doi:10.1175/JAS-D-12-0124.1.

[39] S. A. Young, M. A. Vaughan, A. Garnier, J. L. Tackett, J. D. Lambeth, and K. A. Powell, "Extinction and optical depth retrievals for CALIPSO's Version 4 data release," *Atmos. Meas. Tech.*, vol. 11, pp. 5701-5727, 2018, doi:10.5194/amt-11-5701-2018.

[40] Heymsfield, A. J., D. Winker, and G.-J. van Zadelhoff, "Extinction-ice water content-effective radius algorithms for CALIPSO", *Geophys. Res.Lett.*, vol. 32, doi:10.1029/2005GL022742.

[41] C. Stubenrauch, W. B. Rossow, S. Kinne, S. Ackerman, G. Cesana, H. Chepfer, B. Getzewich, L. DiGirolamo, A. Guignard, A. Heidinger, B. Maddux, P. Menzel, P. Minnis, C. Pearl, S. Platnick, C. Poulsen, J. Riedi, S. Sun-Mack, A. Walther, D. Winker, S. Zeng, and G. Zhao, "Assessment of global cloud datasets from satellites: Project and database initiated by the GEWEX Radiation Panel," *Bull. Am. Meteorol. Soc.*, vol. 94, pp. 1031-1049, 2013, doi:10.1175/BAMS-D-12-00117.

[42] S. Kox, L. Bugliaro, L., and A. Ostler, A., "Retrieval of cloud optical thickness and top altitude from geostationary remote sensing," *Atmos. Meas. Tech.,* vol. 7, pp. 3233-3246, 2014, doi:10.5194/amt-7-3233-2014.

[43] N. Håkansson, C. Adok, A. Thoss, R. Scheirer, and S. Hörnquist, "Neural network cloud top pressure and height for MODIS," *Atmos. Meas. Tech.*, vol. 11, pp. 3177-3196, 2018, doi:10.5194/amt-11-3177-2018.

[44] P. Minnis, S. Sun-Mack, W. L. Smith, Jr., G. Hong, and Y. Chen, "Advances in neural network detection and retrieval of multilayer clouds for CERES using multispectral satellite data," *Proc. SPIE Conf. Remote Sens. Clouds and the Atmos. XXIV*, Strasbourg, France, Sept. 9-12, 11152-1, 12 pp., 2019.

TABLE I. CONTINGENCY TABLE LAYOUT FOR CLOUD PHASE VALIDATION.

cloud phase	CALIOP water	CALIOP ice
CERES-MODIS water	*a*	*b*
CERES-MODIS ice	*c*	*d*

TABLE II. CERES ED4 AQUA MODIS CLOUD PHASE COMPARISON STATISTICS RELATIVE TO CALIOP PHASE CLASSIFICATION FOR SINGLE-PHASE 100% CLOUD-COVERED FOOTPRINTS, HA ≤ 5 KM. JANUARY, APRIL, JULY, & OCTOBER, 2015 & 2016. SIF: SNOW/ICE-FREE, SIC: SNOW/ICE-COVERED

Scene #	Scene Type	Fraction Correct (HR)	Bias	Ice FAR	Water FAR	Hanssen-Kuiper	Number samples	Percent of all matches
	Day							
1	Nonpolar, Land SIF	0.919	-0.049	0.032	0.129	0.846	581	64.7
2	Polar Land SIF	0.928	-0.011	0.097	0.061	0.827	128	61.5
3	Nonpolar Ocean, SIF	0.971	0.006	0.046	0.019	0.941	2,436	71.2
4	Polar Ocean, SIF	0.945	0.023	0.176	0.020	0.871	317	65.0
5	Global, SIF	0.958	-0.003	0.051	0.036	0.911	3,463	69.1
6	Global, SIC	0.919	0.036	0.153	0.037	0.844	769	64.3
	Night							
7	Nonpolar, Land SIF	0.873	0.051	0.137	0.109	0.715	598	68.5
8	Polar Land SIF	0.823	0.132	0.280	0.050	0.679	124	69.3
9	Nonpolar Ocean, SIF	0.918	0.048	0.174	0.027	0.851	2,500	69.5
10	Polar Ocean, SIF	0.840	0.135	0.336	0.023	0.746	384	67.5
11	Global, SIF	0.899	0.061	0.187	0.036	0.817	3,606	69.1
12	Global, SIC	0.798	0.186	0.252	0.034	0.520	1,381	74.0

TABLE III
CERES ED4 AQUA MODIS LIQUID WATER CLOUD-TOP HEIGHT DIFFERENCES (ED4 – CALIOP) RELATIVE TO CALIOP (HA ≤ 5) FOR OVERCAST FOOTPRINTS. JANUARY, APRIL, JULY, & OCTOBER, 2010. SIF: SNOW/ICE-FREE, SIC: SNOW/ICE-COVERED

Category	**A**: Single Layer only			**B**: Single-layer and multi-layer, liquid only			**C**: All with liquid top layer		
Parameter	Mean CTH_{CA} (km)	Bias (SDD) (km)	% of all	Mean CTH_{CA} (km)	Bias (SDD) (km)	% of all	Mean CTH_{CA} (km)	Bias (SDD) (km)	# Samples x 10^{-3}
Day									
Land, SIF	3.13	-0.10 (0.91)	51	3.52	-0.31 (1.05)	78	3.81	-0.53 (1.31)	223
Ocean, SIF	1.87	-0.00 (0.67)	69	2.08	-0.12 (0.82)	86	2.39	-0.34 (1.19)	1,092
Global, SIC	1.98	0.21 (1.00)	60	2.09	0.08 (1.08)	79	2.27	-0.10 (1.31)	330
Global, All	2.03	0.03 (0.77)	65	2.27	-0.11 (0.91)	84	2.56	-0.32 (1.24)	1,646
Night									
Land, SIF	3.15	0.09 (0.87)	61	3.49	-0.08 (0.99)	88	3.69	-0.24 (1.18)	140
Ocean, SIF	1.79	0.15 (0.63)	75	1.97	0.05 (0.74)	91	2.19	-0.10 (1.04)	1,057
Global, SIC	1.61	0.11 (0.90)	68	1.74	-0.02 (0.98)	90	1.87	-0.14 (1.14)	202
Global, All	1.88	0.14 (0.70)	72	2.08	0.03 (0.81)	91	2.29	-0.12 (1.07)	1,399

TABLE IV
CERES ED4 AQUA MODIS ICE CLOUD HEIGHT DIFFERENCES (ED4 – CALIOP) FOR SINGLE-PHASE ED4 ICE, OPAQUE CALIOP PIXELS 100% CLOUD-COVERED FOOTPRINTS. JANUARY, APRIL, JULY, & OCTOBER, 2010. SIF: SNOW/ICE-FREE, SIC: SNOW/ICE-COVERED

Category	**A**: Single Layer only			**B**: Single layer and multi-layer, ice only			**C:** All with ice top layer		
Parameter	Mean CTH_{CA} (km)	Bias (SDD) (km)	% of all	Mean CTH_{CA} (km)	Bias (SDD) (km)	% of all	Mean CTH_{CA} (km)	Bias (SDD) (km)	# Samples x 10^3
Day									
Global, Effective Height	11.85	-2.31 (1.27)	57	11.20	-2.32 (1.38)	72	10.98	-2.69 (1.58)	746
Global, Archived Top	11.85	-1.75 (1.36)	57	11.20	-1.65 (1.50)	72	10.98	-1.90 (1.68)	746
Global, Corrected Top	11.85	-0.74 (1.29)	57	11.20	-0.81 (1.35)	72	10.98	-1.18 (1.58)	746
SIF Land, Corrected	12.12	-0.52 (1.28)	58	11.73	-0.62 (1.36)	77	11.57	-0.99 (1.61)	148
SIF Ocean, Corrected	12.15	-0.77 (1.24)	58	11.57	-0.83 (1.30)	73	11.35	-1.21 (1.55)	491
SIC Global, Corrected	9.28	-1.08 (1.51)	40	8.49	-0.98 (1.53)	65	8.47	-1.32 (1.69)	107
Night									
Global, Effective Height	11.50	-2.03 (1.65)	50	10.98	-2.08 (1.90)	61	10.94	-2.67 (2.29)	790
Global, Archived Top	11.50	-1.37 (1.66)	50	10.98	-1.33 (1.94)	61	10.94	-1.70 (2.19)	790
Global, Corrected Top	11.50	-0.71 (1.64)	50	10.98	-0.78 (1.91)	61	10.94	-1.36 (2.24)	790
SIF Land, Corrected	12.72	-0.54 (1.50)	53	12.41	-0.58 (1.53)	60	12.58	-1.33 (2.15)	149
SIF Ocean, Corrected	11.74	-0.49 (1.42)	50	11.16	-0.51 (1.46)	61	11.04	-1.12 (1.88)	451
SIC Global, Corrected	9.74	-1.43 (2.04)	45	9.48	-1.54 (2.71)	62	9.40	-1.93 (2.90)	190

TABLE V
CERES ED4 AQUA MODIS ICE CLOUD HEIGHT DIFFERENCES (ED4 – CALIOP) FOR SINGLE-PHASE ED4 ICE, NON-OPAQUE CALIOP PIXELS 100% CLOUD-COVERED FOOTPRINTS. JANUARY, APRIL, JULY, & OCTOBER, 2010. SIF: SNOW/ICE-FREE, SIC: SNOW/ICE-COVERED

Category	**A**: Single Layer only			**B**: Single layer and multi-layer, ice only			**C:** All with ice top layer		
Parameter	Mean CTH_{CA} (km)	Bias (SDD) [CEH_{Ed4}-CTH_{CA}] (km)	% of all	Mean CTH_{CA} (km)	Bias (SDD) [CEH_{Ed4}-CTH_{CA}] (km)	% of all	Mean CTH_{CA} (km)	Bias (SDD) [CEH_{Ed4}-CTH_{CA}] (km)	# Samples x 10^3
Day									
SIF Land	12.07	-1.21 (2.09) [-2.53]	61	11.72	-1.23 (2.15) [-2.56]	75	11.78	-1.37 (2.22) [-2.69]	62
SIF Ocean	12.96	-2.11 (2.73) [-3.42]	60	12.69	-2.04 (2.75) [-3.30]	75	12.62	-2.01 (2.72) [-3.27]	189
Global SIC	8.52	-2.21 (2.47) [-3.65]	42	7.76	-1.84 (2.37) [-3.37]	78	7.73	-1.92 (2.30) [-3.46]	88
Global, All	11.92	-1.95 (2.59) [-3.29]	57	11.19	-1.84 (2.57) [-3.19]	76	11.20	-1.87 (2.54) [-3.22]	339
Night									
SIF Land	12.57	-0.20 (2.03) [-0.87]	57	12.21	-0.20 (2.07) [-0.92]	68	12.47	-0.58 (2.27) [-1.40]	160
SIF Ocean	12.85	-0.36 (1.82) [-1.36]	43	12.40	-0.37 (1.91) [-1.41]	51	12.49	-0.88 (2.20) [-2.08]	366
Global SIC	9.29	-1.85 (3.24) [-2.89]	46	8.84	-1.78 (3.73) [-2.91]	75	8.89	-1.97 (3.79) [-3.12]	408
Global, All	11.26	-0.97 (2.68) [-1.92]	47	10.55	-1.06 (3.09) [-2.08]	64	10.91	-1.31 (3.07) [-2.42]	935

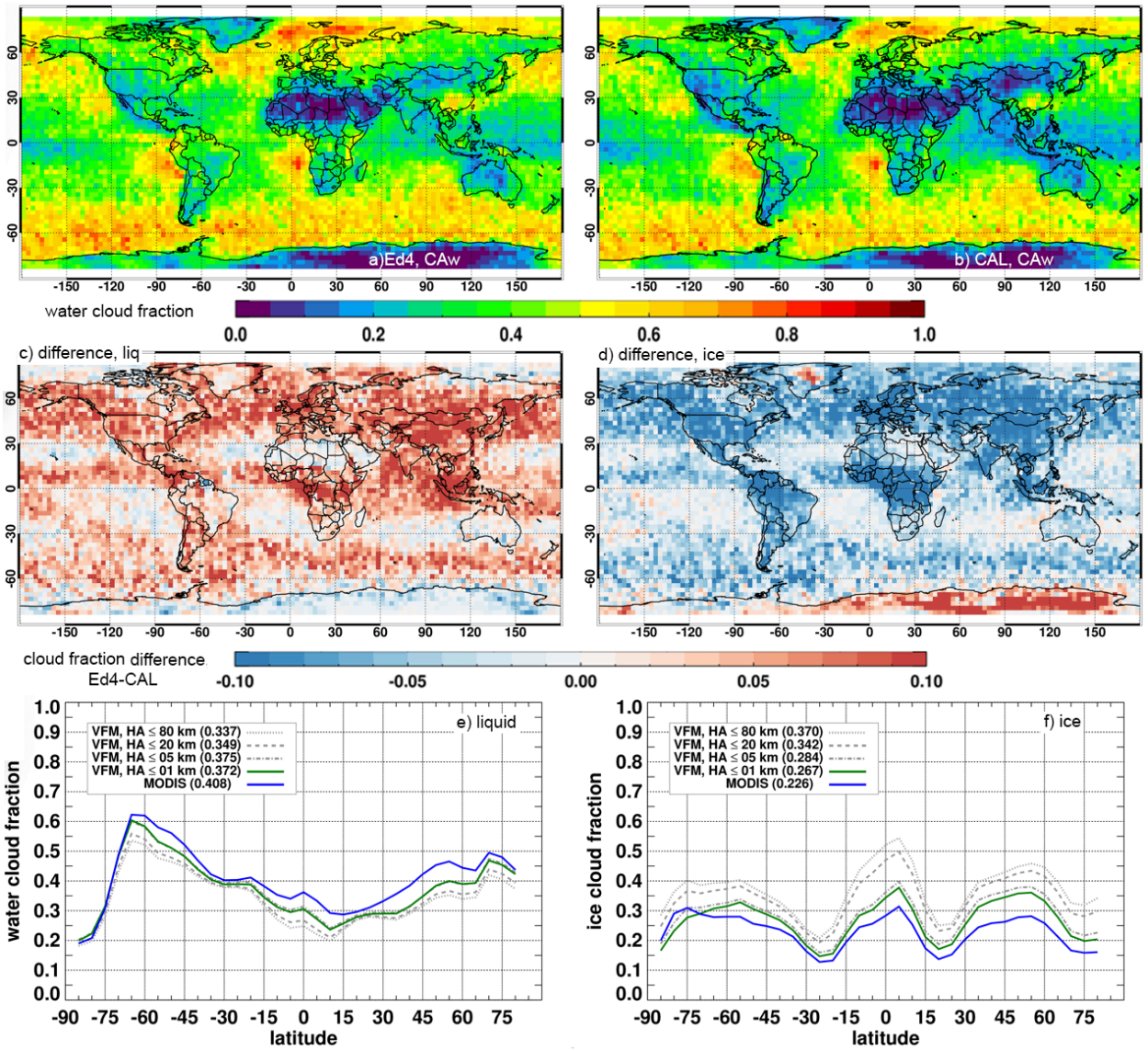

Fig. 1. Daytime regional and zonal comparisons of Ed4 Aqua MODIS and CALIOP cloud phase fractions, and their differences for JAJO, 2015-2016. Mean 3° liquid cloud fractions from Ed4 (a) and from CALIOP for HA $\leq$ 5 km (b). Regional mean cloud fraction difference, Ed4 – CALIOP, for liquid (c) and ice (d). Zonal mean cloud fraction differences, Ed4 – CALIOP, for liquid (e) and ice (f) for various HA resolutions.

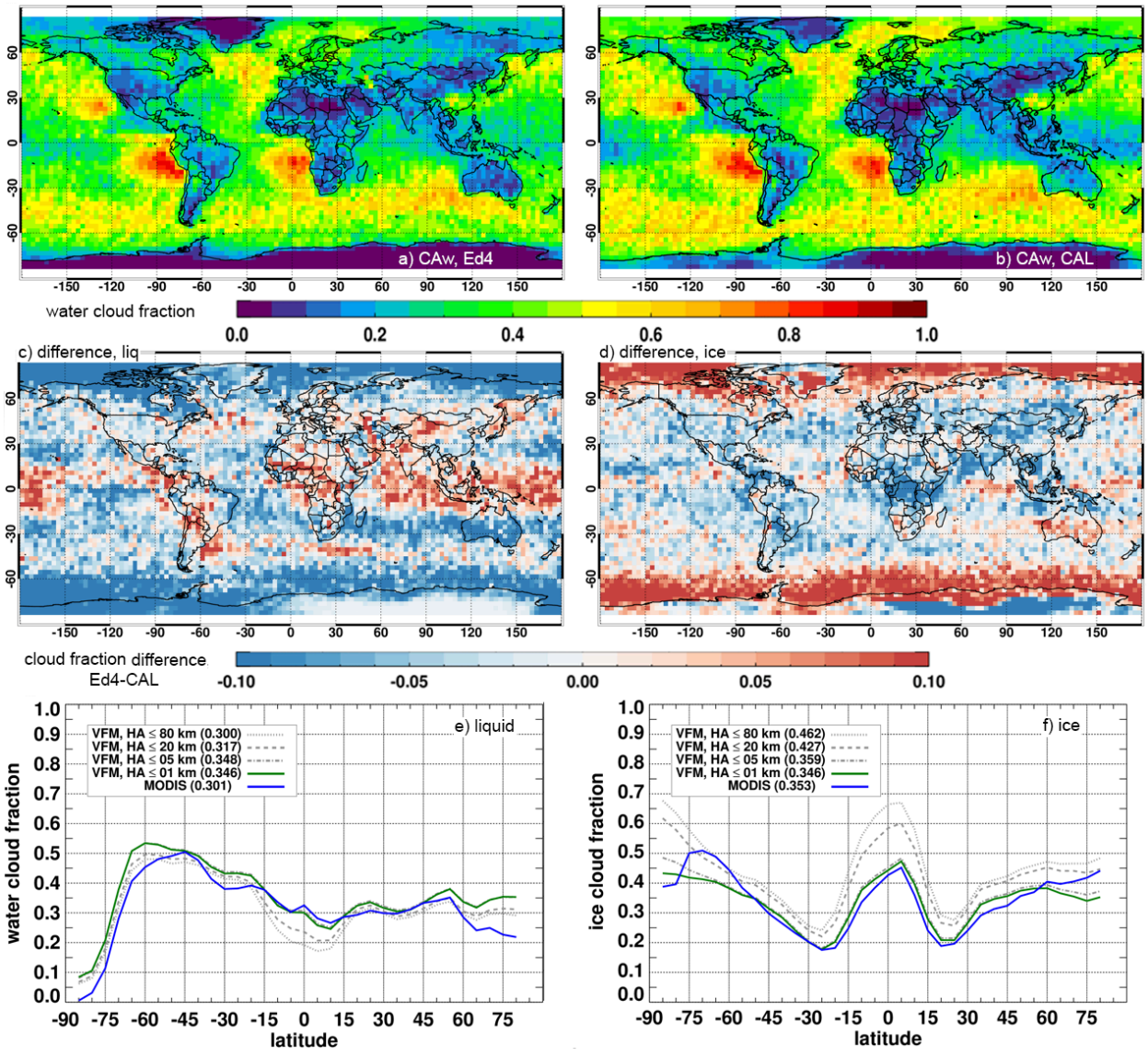

Fig. 2. Same as Fig. 1, except for night.

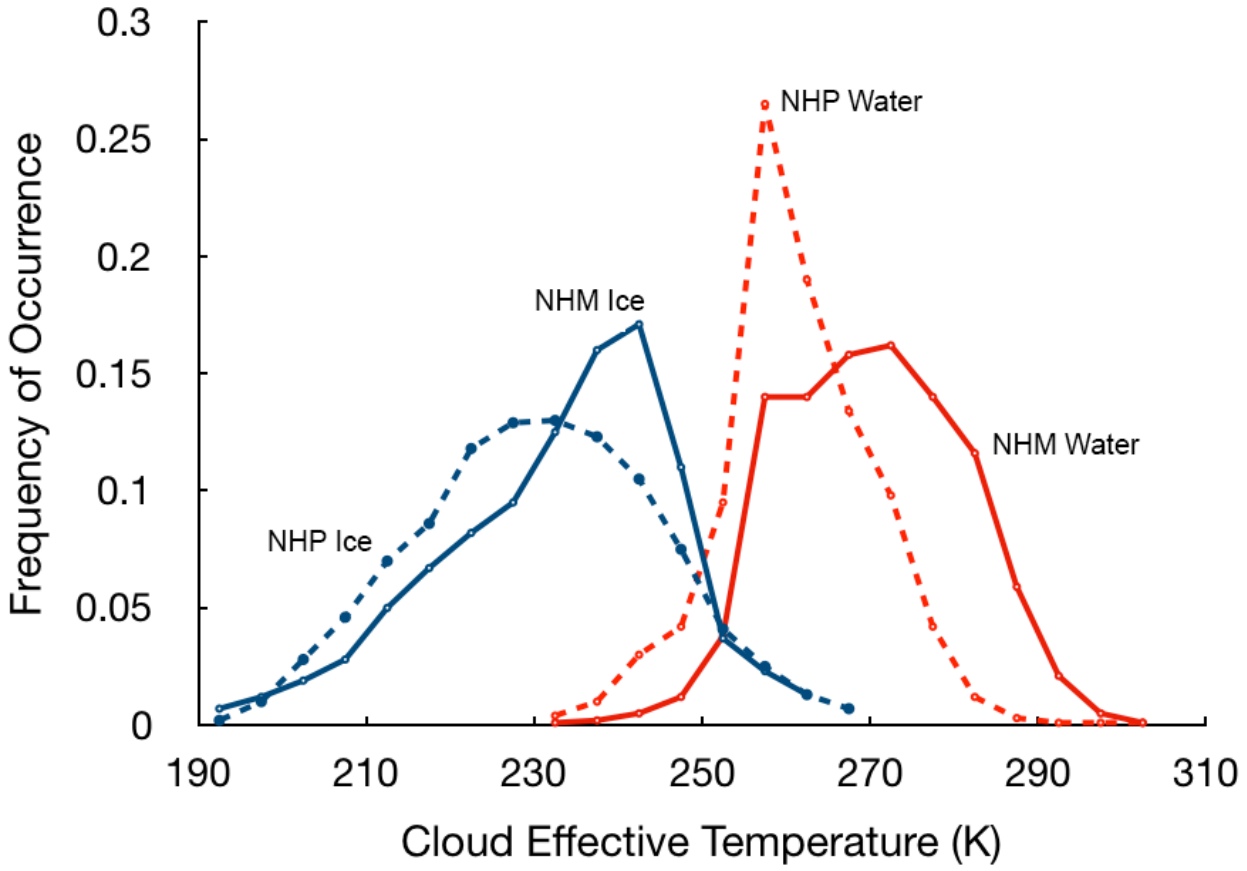

Fig. 3. Histogram of nocturnal cloud phase for 2016 Aqua Ed4 northern polar (NHP) and midlatitude (NHM) zones as a function of cloud effective temperature.

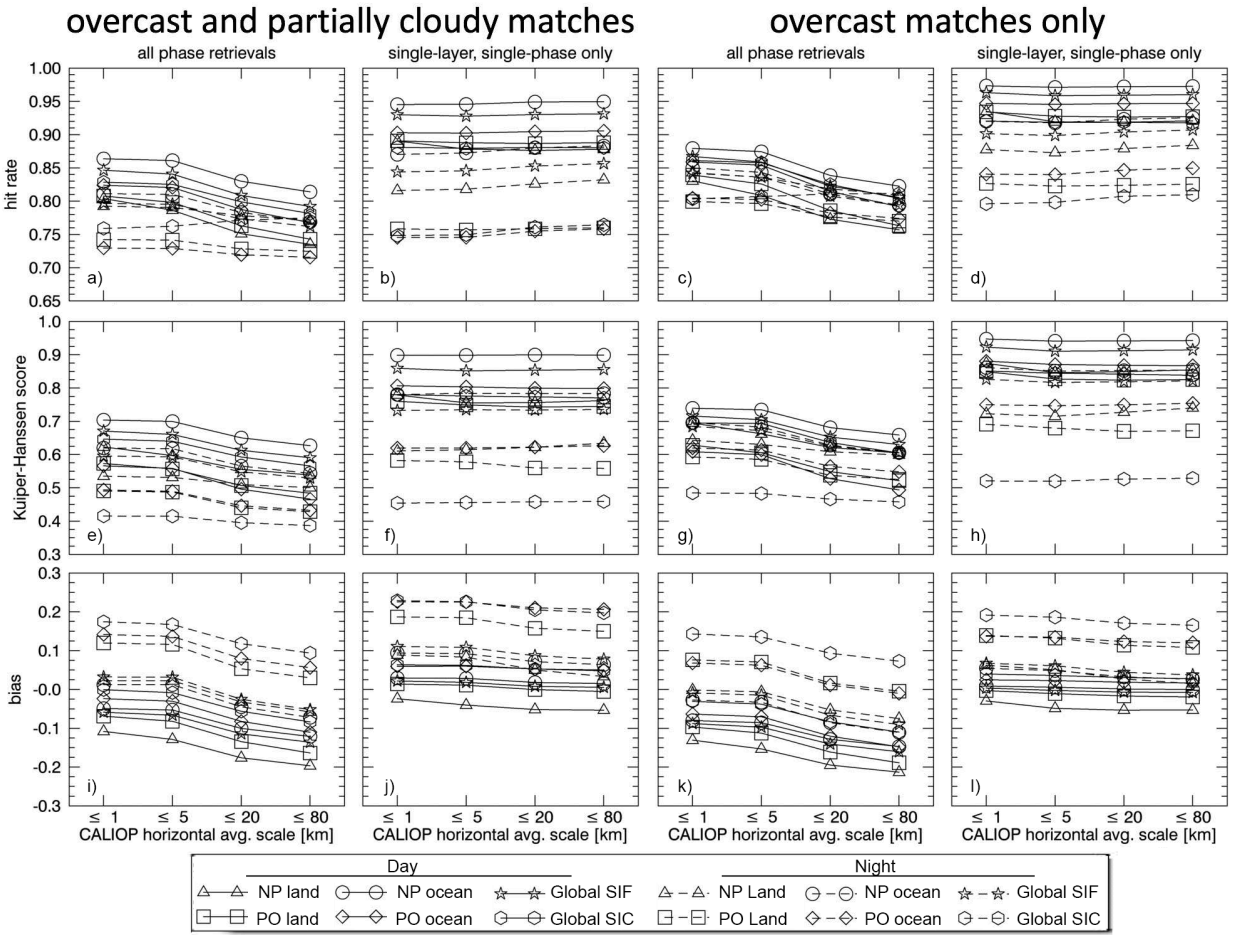

Fig. 4. Cloud phase hit rate, Hansson-Kuiper skill score, and bias plotted for different surface conditions for daytime and nighttime overpasses as functions of CALIOP HA.

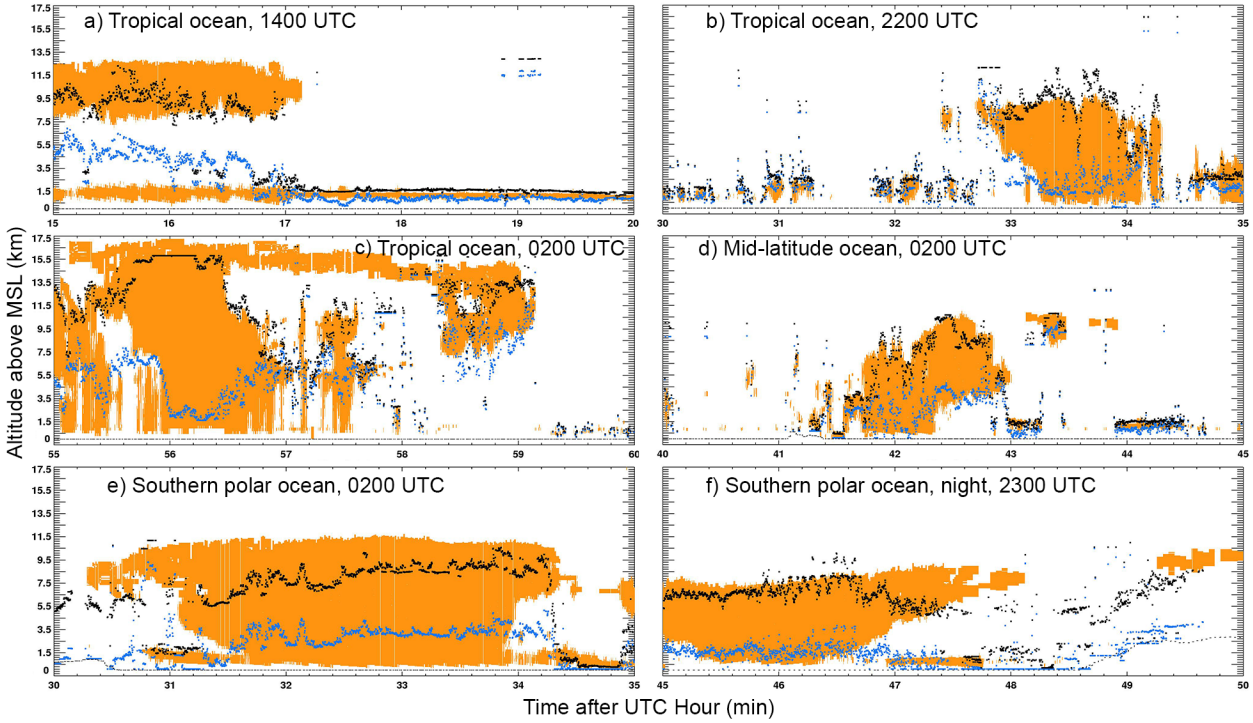

Fig. 5. Cloud boundaries from C3M active sensor data (orange) and CERES Ed4 Aqua retrievals, 2 October, 2009. CERES cloud base and top are shown in blue and black, respectively. All profiles are for daytime unless indicated otherwise.

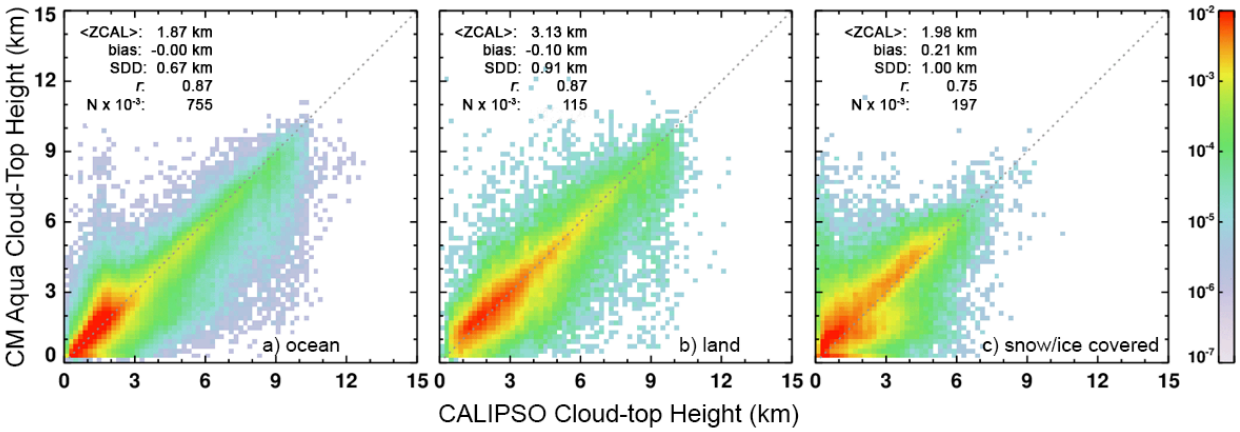

Fig. 6. Comparison of cloud top heights for overcast single-layer liquid clouds identified by both Ed4 and CALIOP for (a) ice-free ocean, (b) snow-free land, and for (c) snow/ice-covered surfaces, observed during daytime in January, April, July, and October 2010. Only scenes with single-layer liquid topped clouds, as determined from CALIOP, are included.

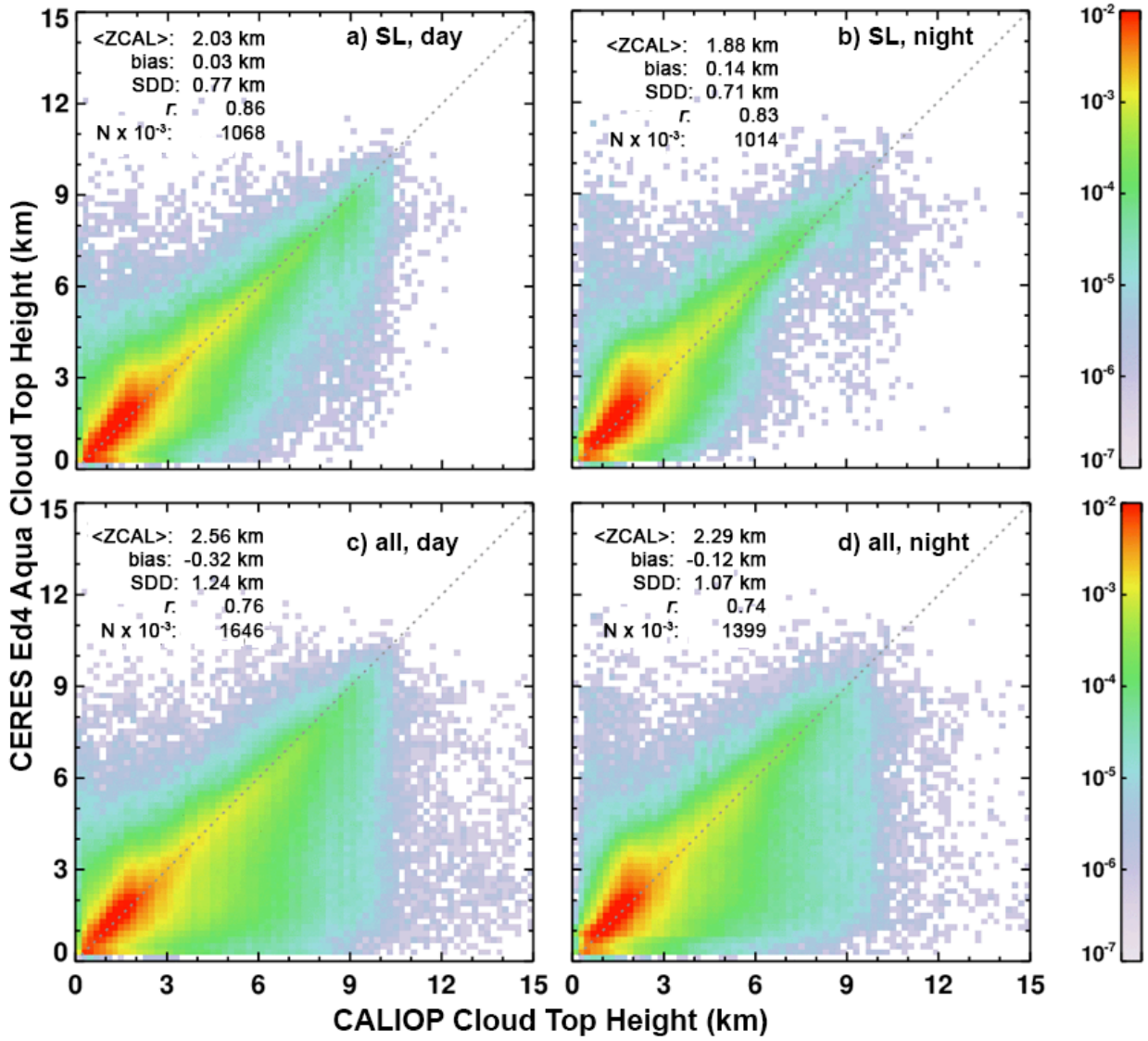

Fig. 7. Scatterplots of CALIOP and Ed4 water-phase cloud top heights under the following conditions; (a) daytime, single-layer only, (b) nighttime, single-layer, (c) daytime, all liquid topped single-multi-layer clouds, and (d) nighttime, all liquid topped single-layer and multi-layer clouds. All scenes are overcast.

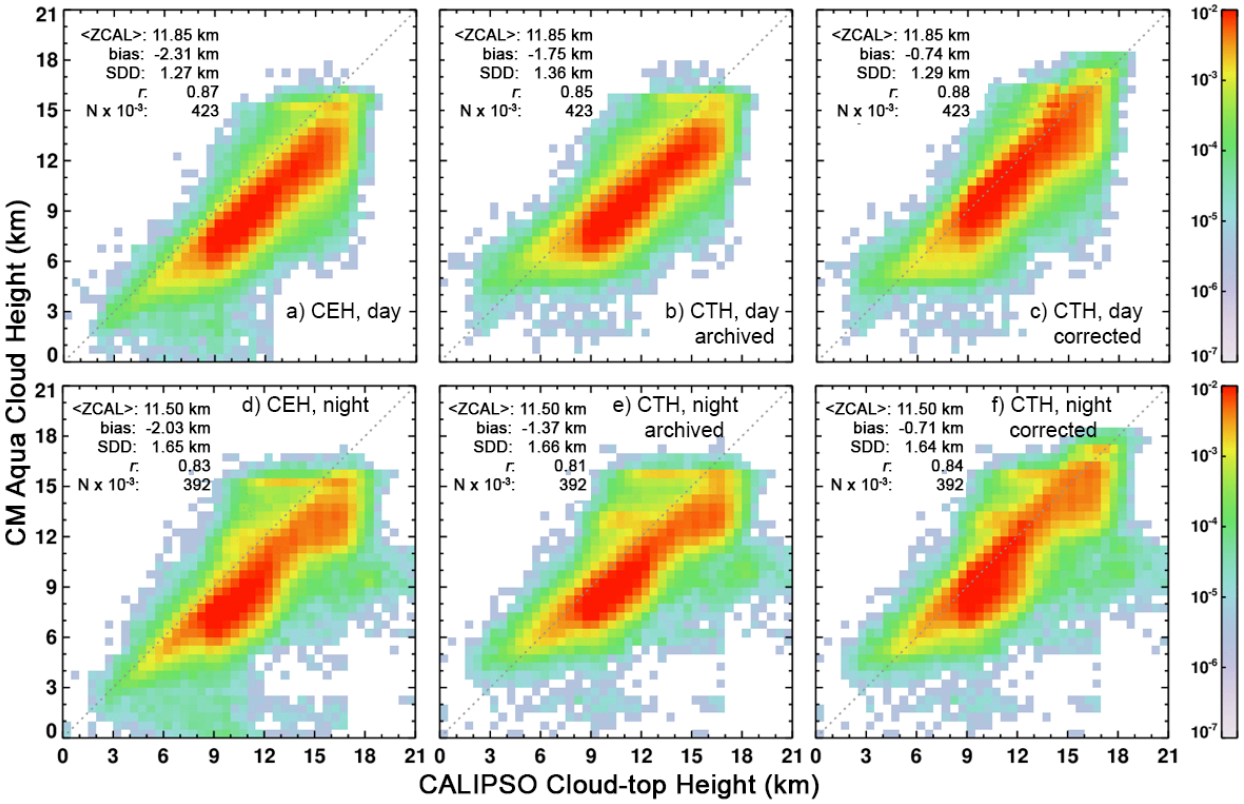

Fig. 8. Comparison of matched overcast CALIOP opaque single-layer ice cloud tops and Ed4 Aqua cloud height parameters, cloud effective height (left), archived cloud top height (middle), and *ex post facto* corrected cloud top height (right), for all surface types during daytime (top) and night (bottom) in January, April, July, and October 2010. Opaque is defined by the lack of a surface return in the CALIOP signal.

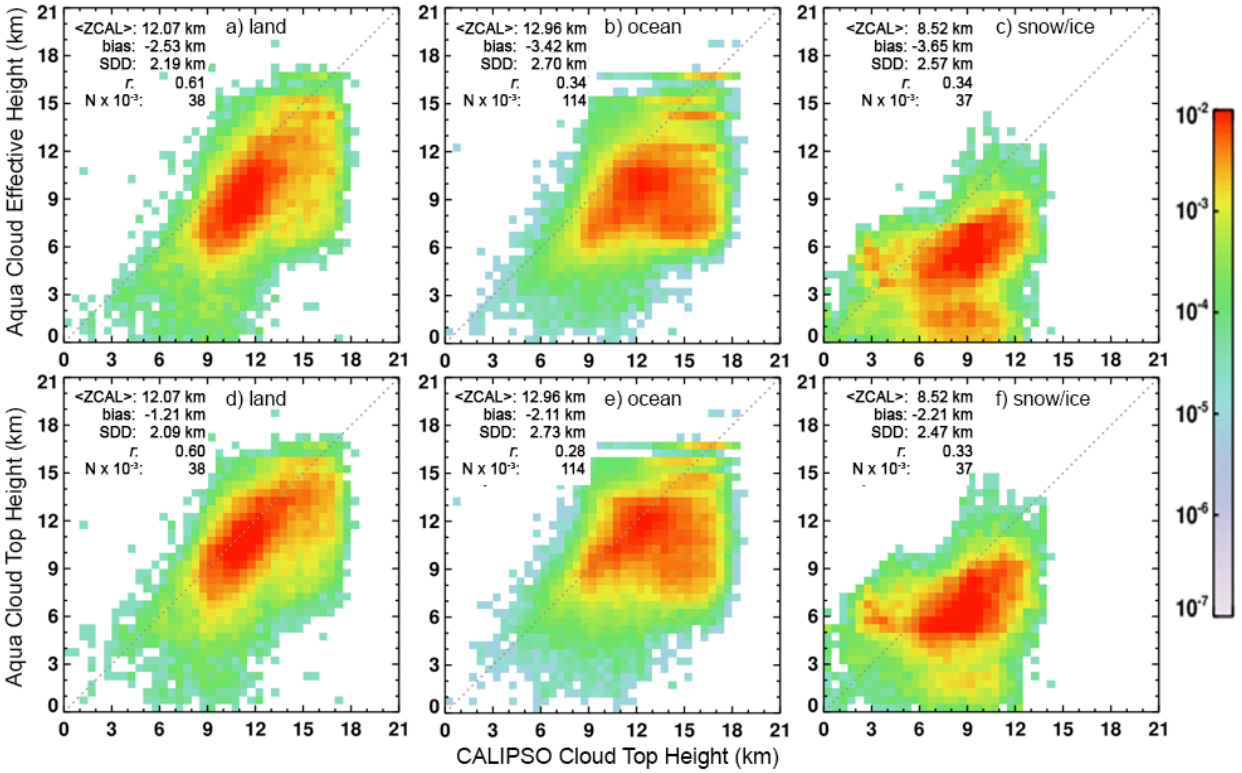

Fig. 9. Comparison of matched overcast CALIOP single-layer, non-opaque ice cloud tops and Ed4 Aqua cloud height parameters, cloud effective height (top) and cloud top height (bottom), during daytime for snow-free land (left) and ocean (middle), and snow-covered surfaces (right) observed during January, April, July, and October 2010. <ZCAL> indicates average of CALIOP cloud top heights. Non-opaque is defined by the presence of a surface return in the CALIOP signal. CTH_{CM} data include heights that have been corrected *ex post facto* with the Ed4 thick cloud top adjustment.

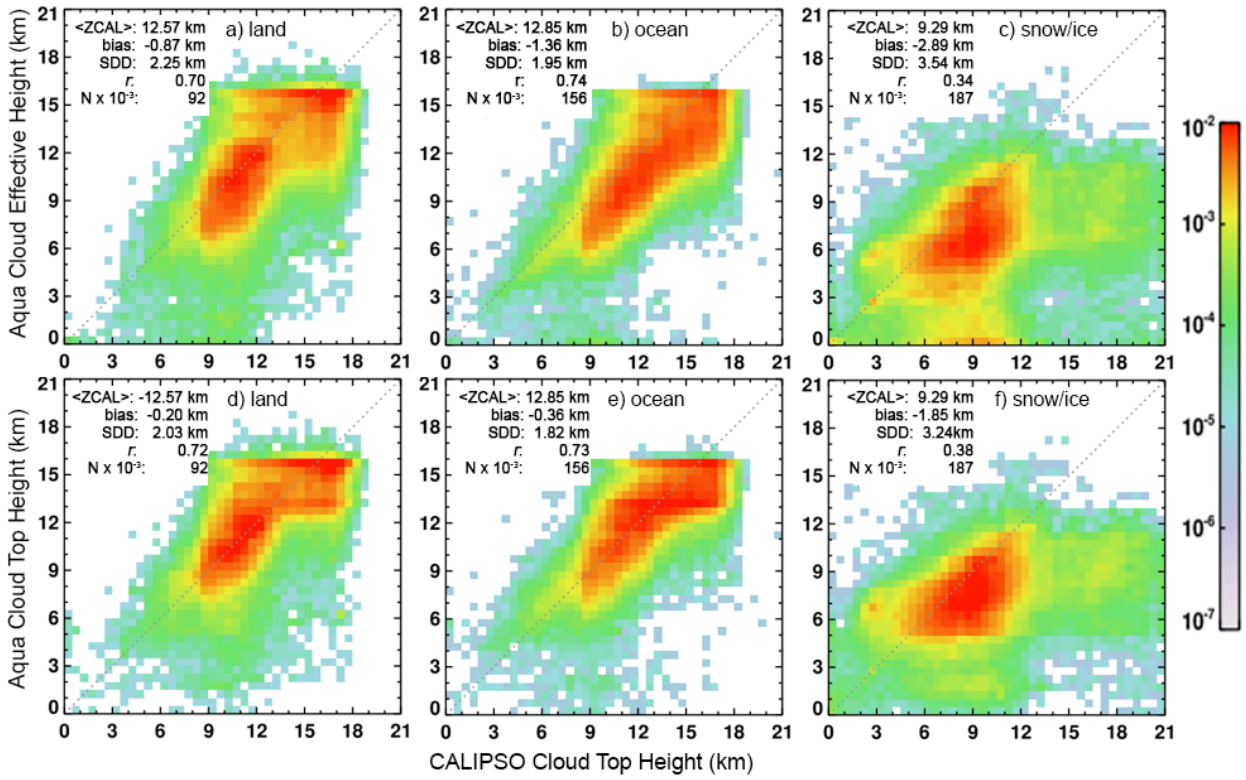

Fig. 10. Same as Fig. 9, except for night.

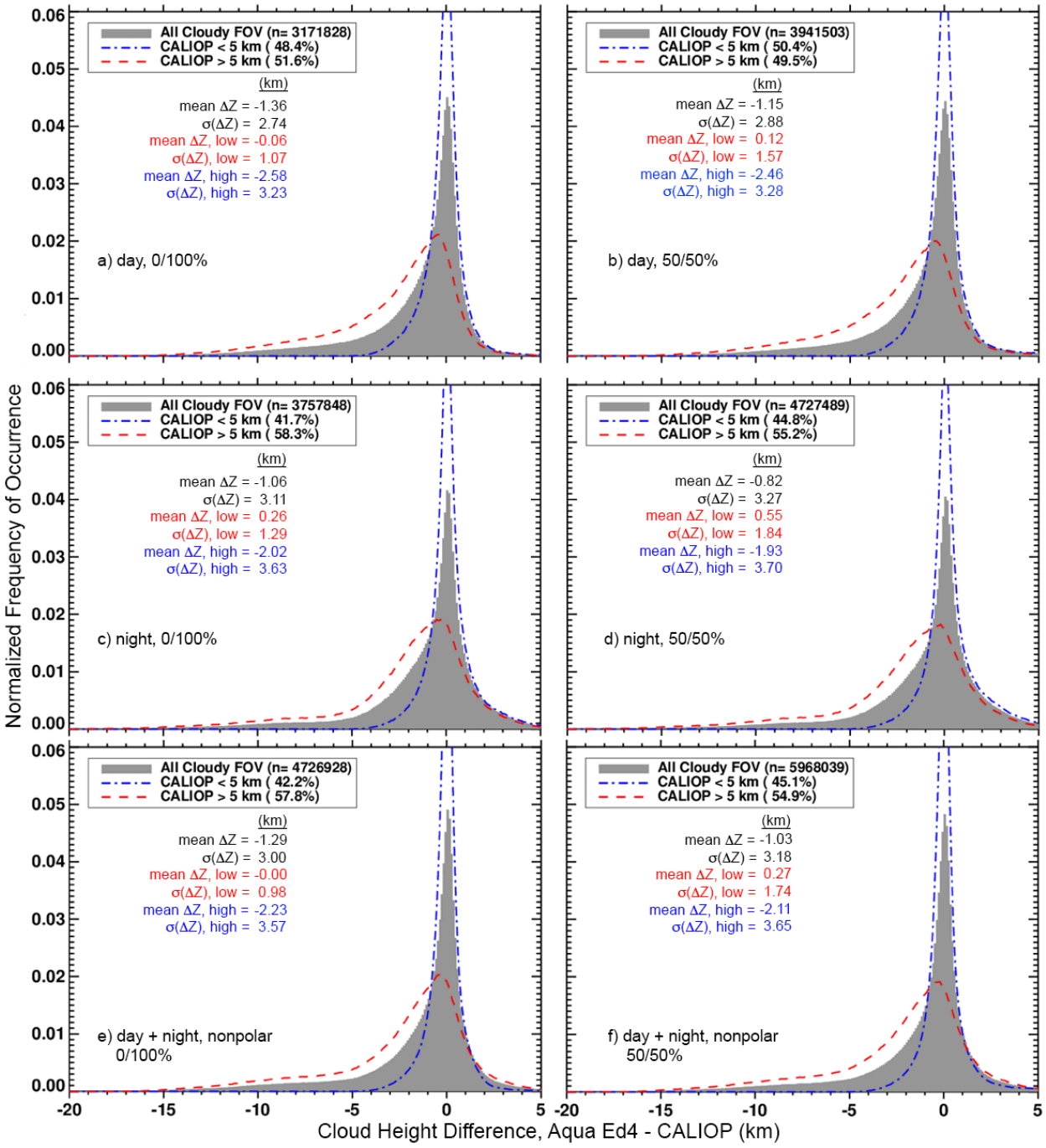

Fig. 11. Probability distribution functions of cloud top height differences between CERES Aqua and matched CALIOP cloud pixels, January, April, July, and October 2010. Left: CALIOP pixels are overcast. Right: CALIPSO pixel is cloudy if cloud amount exceeds 50%. Top: day. Middle: night. Bottom: Nonpolar, day + night. Low and high clouds defined by CALIPSO cloud top heights $\leq$ 5 km and > 5 km, respectively.

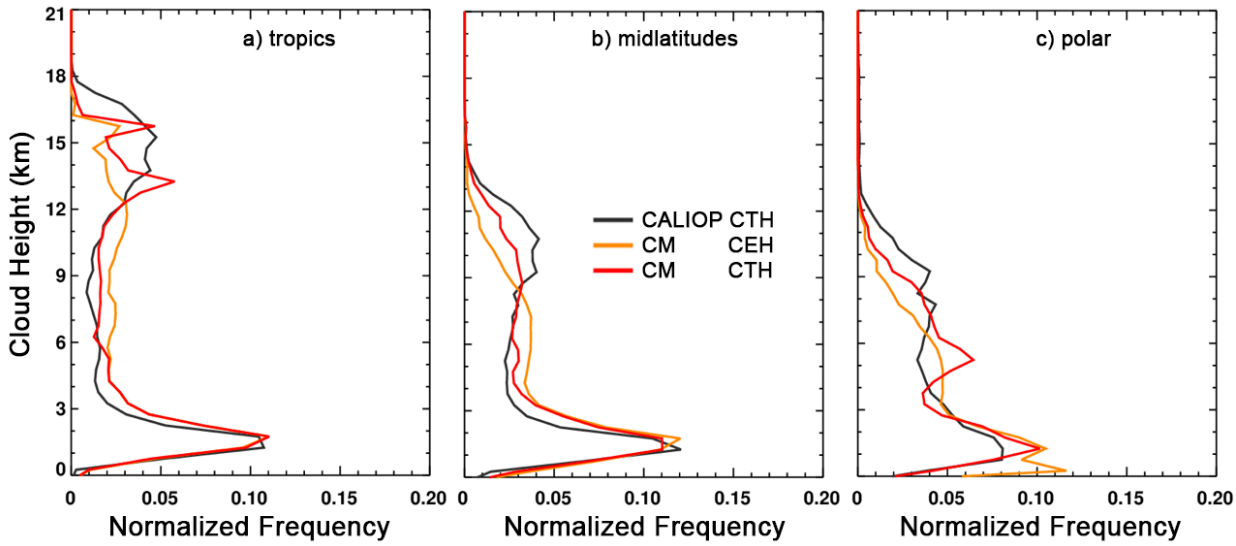

Fig. 12. Vertical distribution of cloud heights from CALIOP and CERES Ed4 Aqua MODIS from JAJO 2010. Black, red and orange lines correspond to CALIOP topmost cloud-top height, CERES cloud-top height, and CERES cloud effective height, respectively. *Ex post facto* correction applied to opaque ice clouds.

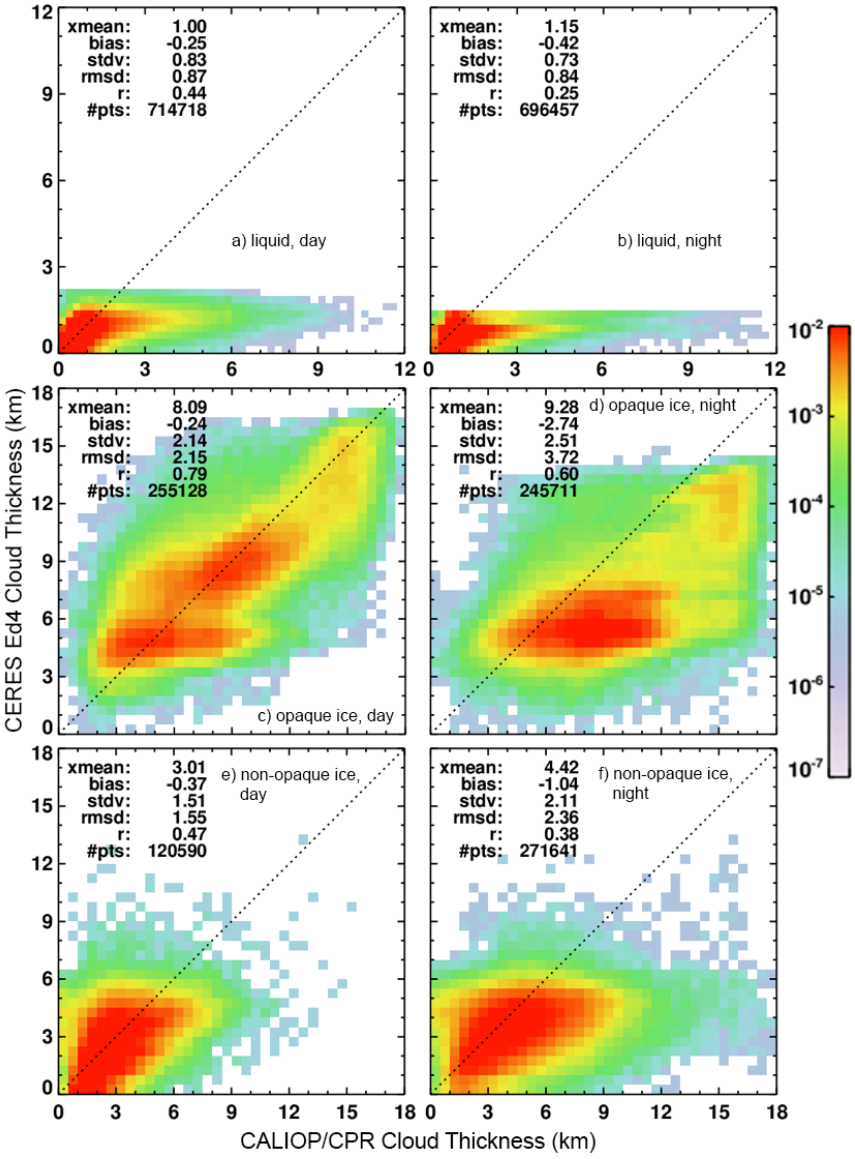

Fig. 13. Single-layer cloud thickness comparisons for water (top row), opaque ice (middle row) and non-opaque ice (bottom) clouds for JAJO 2010. Left: daytime, Right: night.

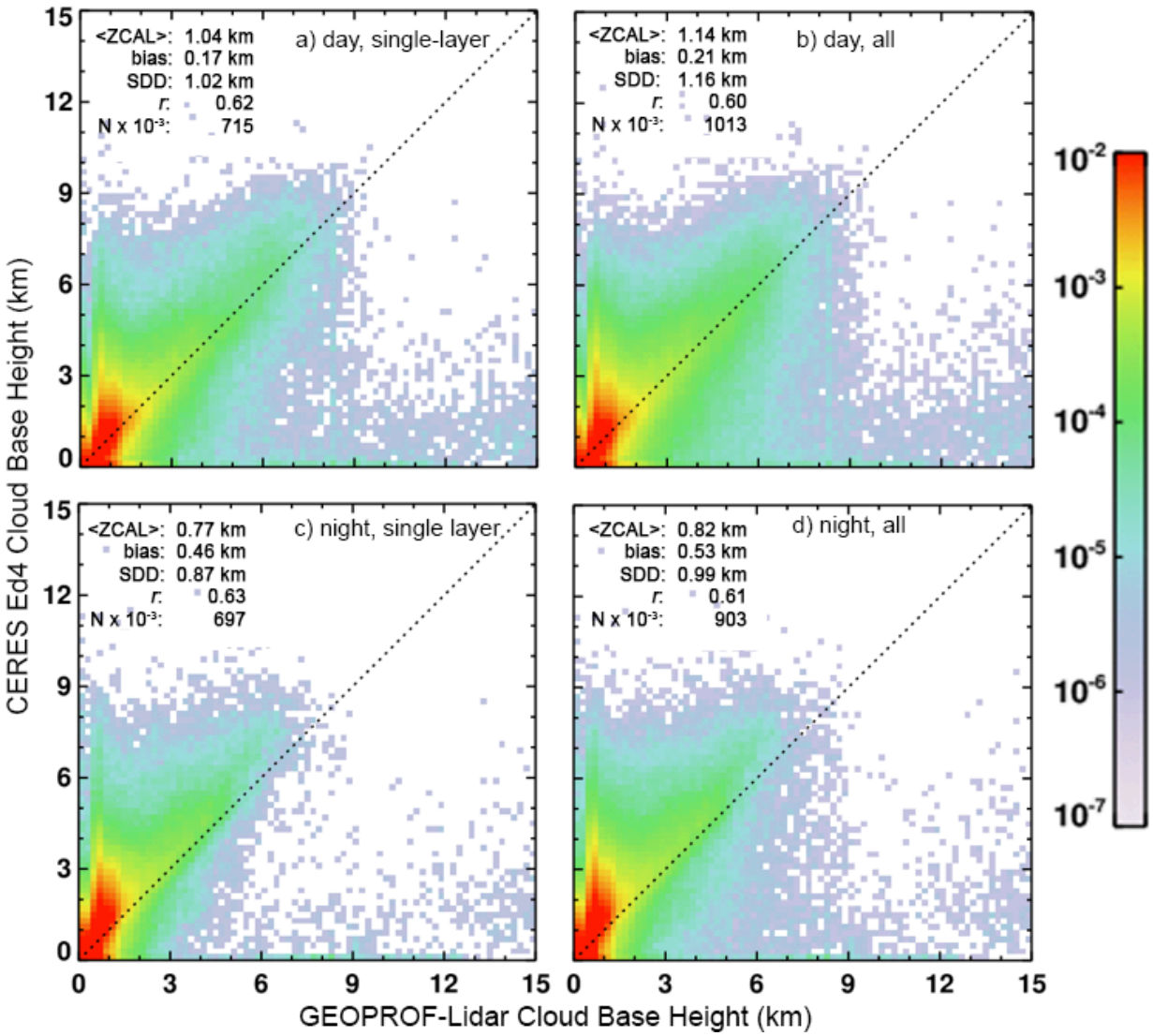

Fig. 14. Cloud base height comparison for CloudSat-CALIPSO GEOPPROF-Lidar product and overcast CERES Ed4 Aqua MODIS liquid water clouds, JAJO 2010. Top row: daytime, bottom row: night. Left: GEOPROF-Lidar single-layer liquid water only, right: all GEOPROF-Lidar liquid-water topmost cloud layer.

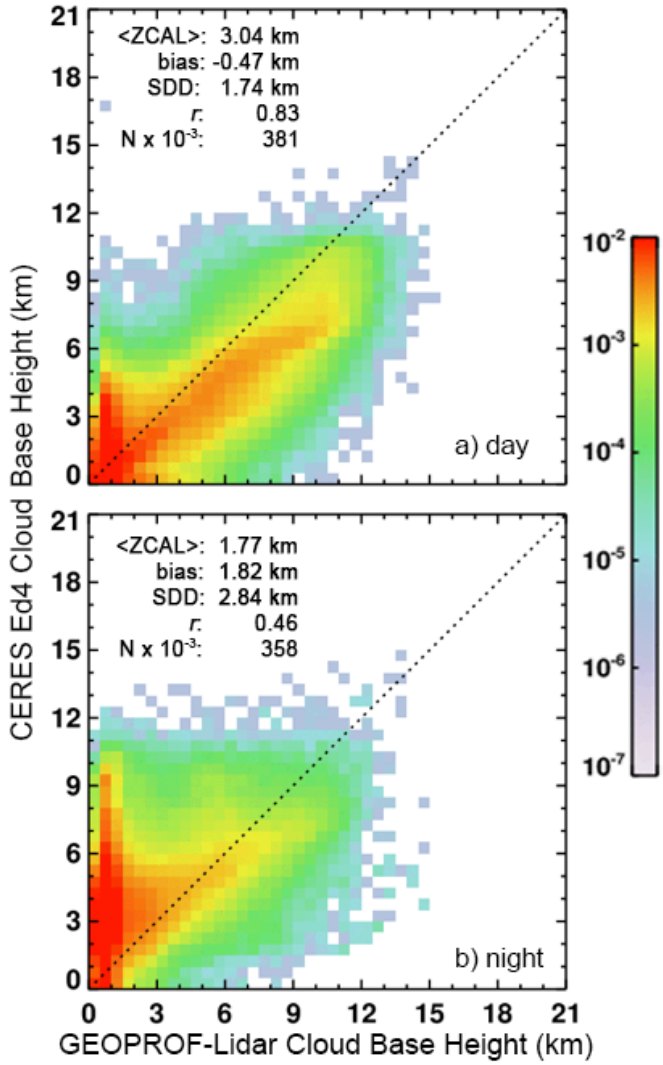

Fig. 15. Comparison of cloud base heights from matched overcast CERES Aqua MODIS ice clouds and GEOPROF-Lidar opaque ice-topped clouds, JAJO 2010.

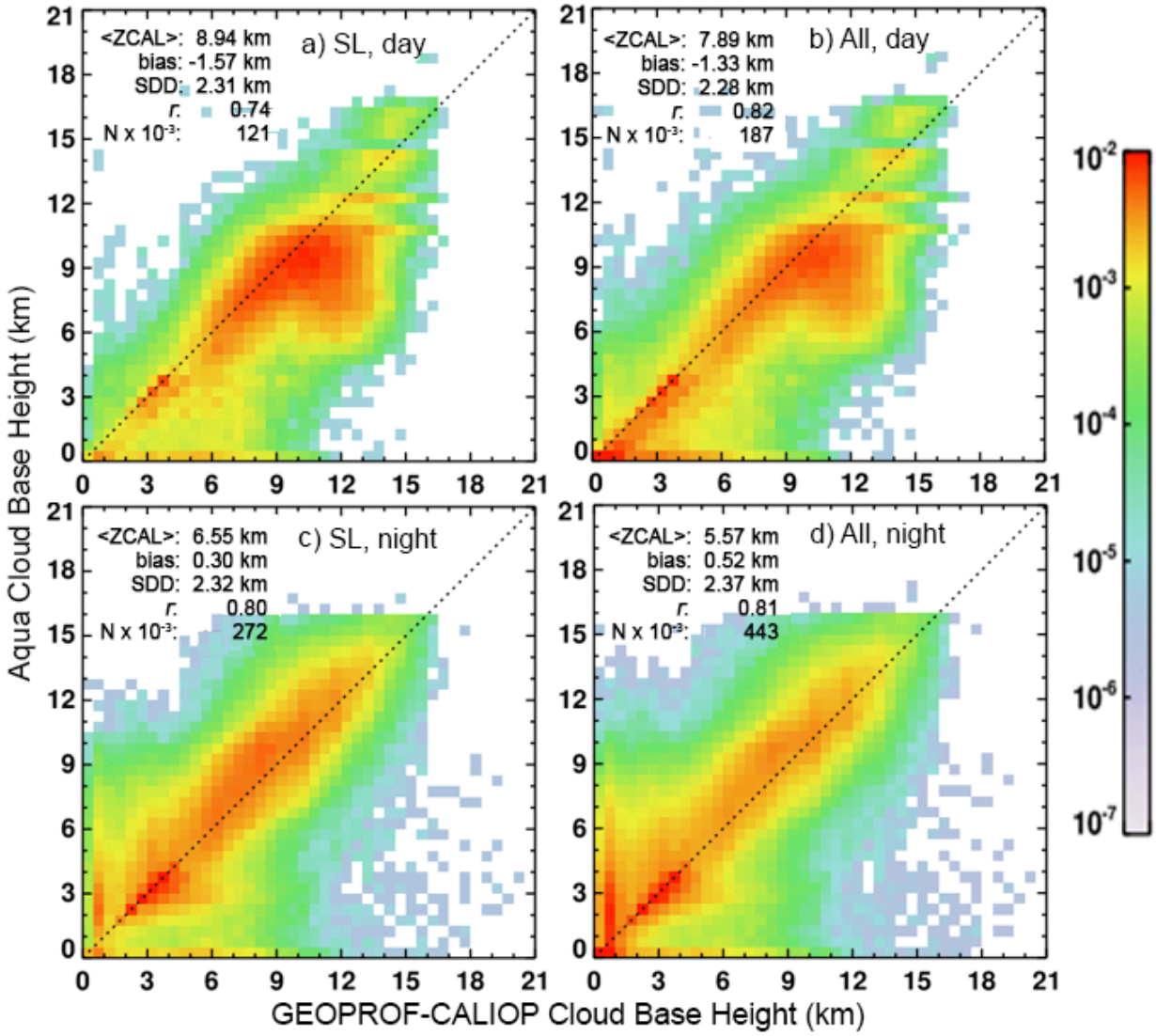

Fig. 16. Comparison of matched overcast, RL-GEOPROF and Aqua Ed4 cloud bases for single-layer, non-opaque ice clouds (top) and all CALIOP non-opaque ice clouds (bottom) over all surfaces during daytime (left) and night (right) observed during January, April, July, and October 2010. Non-opaque is defined by the presence of a surface return in the CALIOP signal for single-layer clouds. Non-opaque for all conditions requires that either the surface or a low-level water cloud is detected underneath the ice clouds. See text for definition of all conditions.

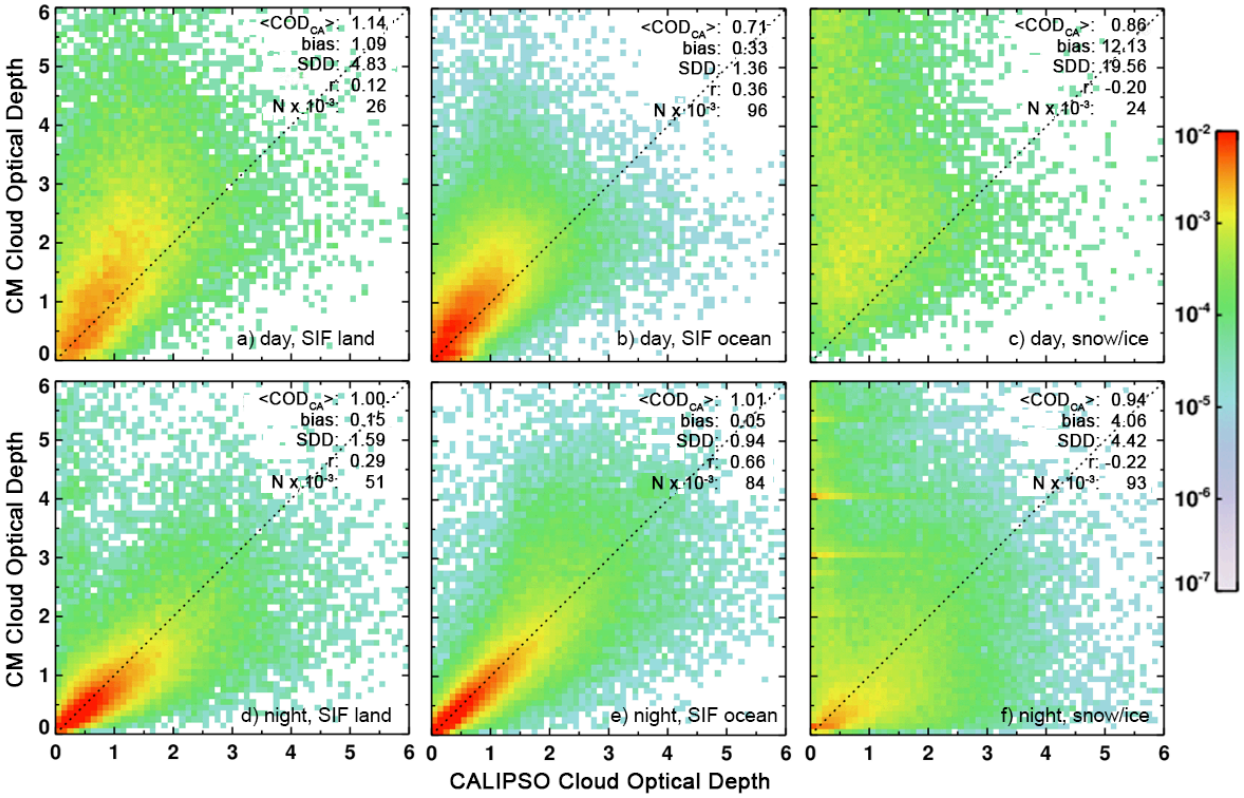

Fig. 17. Comparison of CERES Ed4 Aqua MODIS cloud optical depths with those derived from CALIOP backscatter using the Version 4 analysis method, JAJO 2010. Top: day, bottom: night. Left: snow-free land, middle: snow-free ocean, right: ice/snow-covered surfaces.

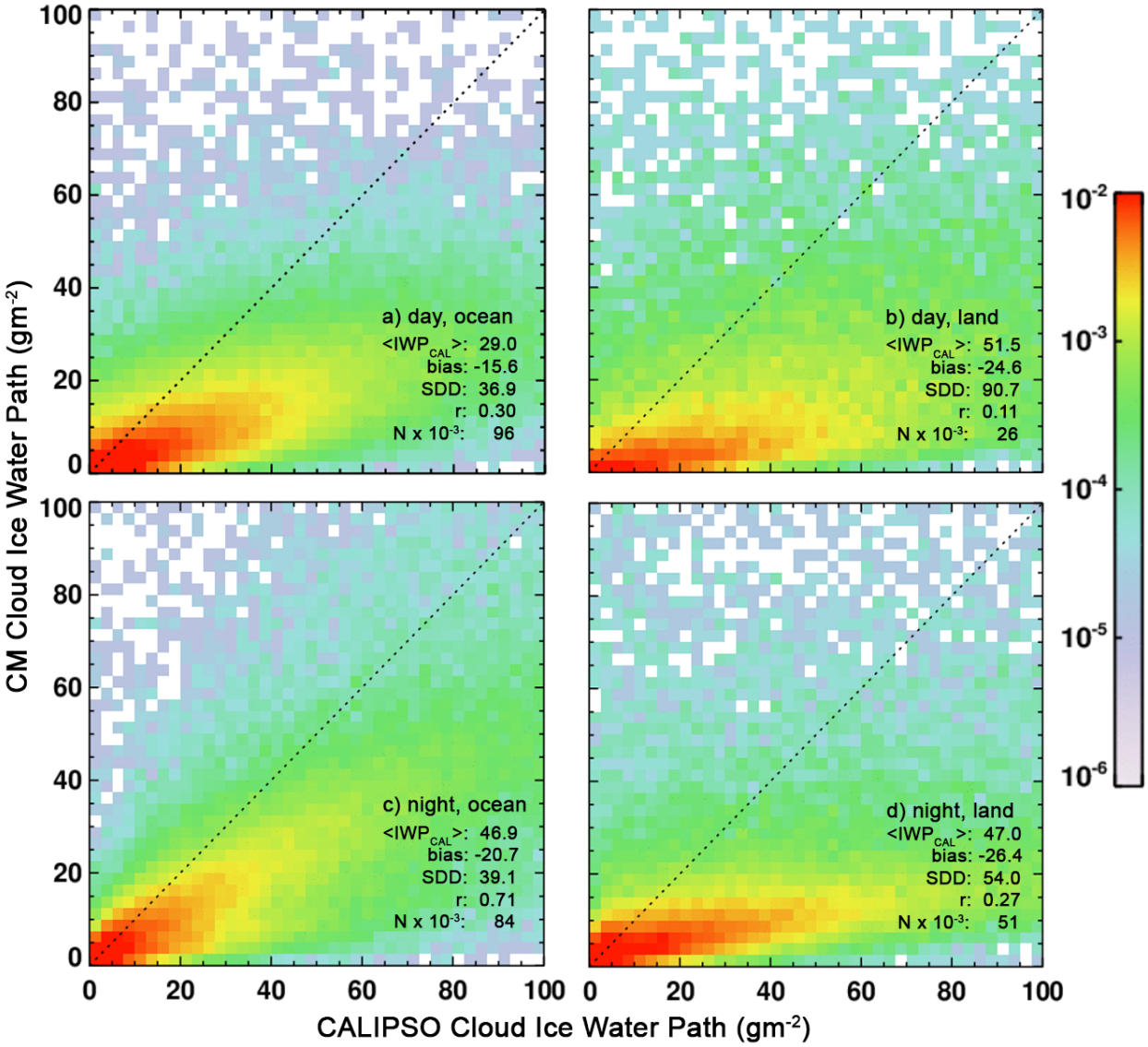

Fig. 18. Comparison of Aqua MODIS Ed4 cloud ice water paths with those derived from CALIOP backscatter using the Version 4 constrained method, JAJO 2010. Top: day, bottom: night. Left: snow-free ocean, right: snow-free land.

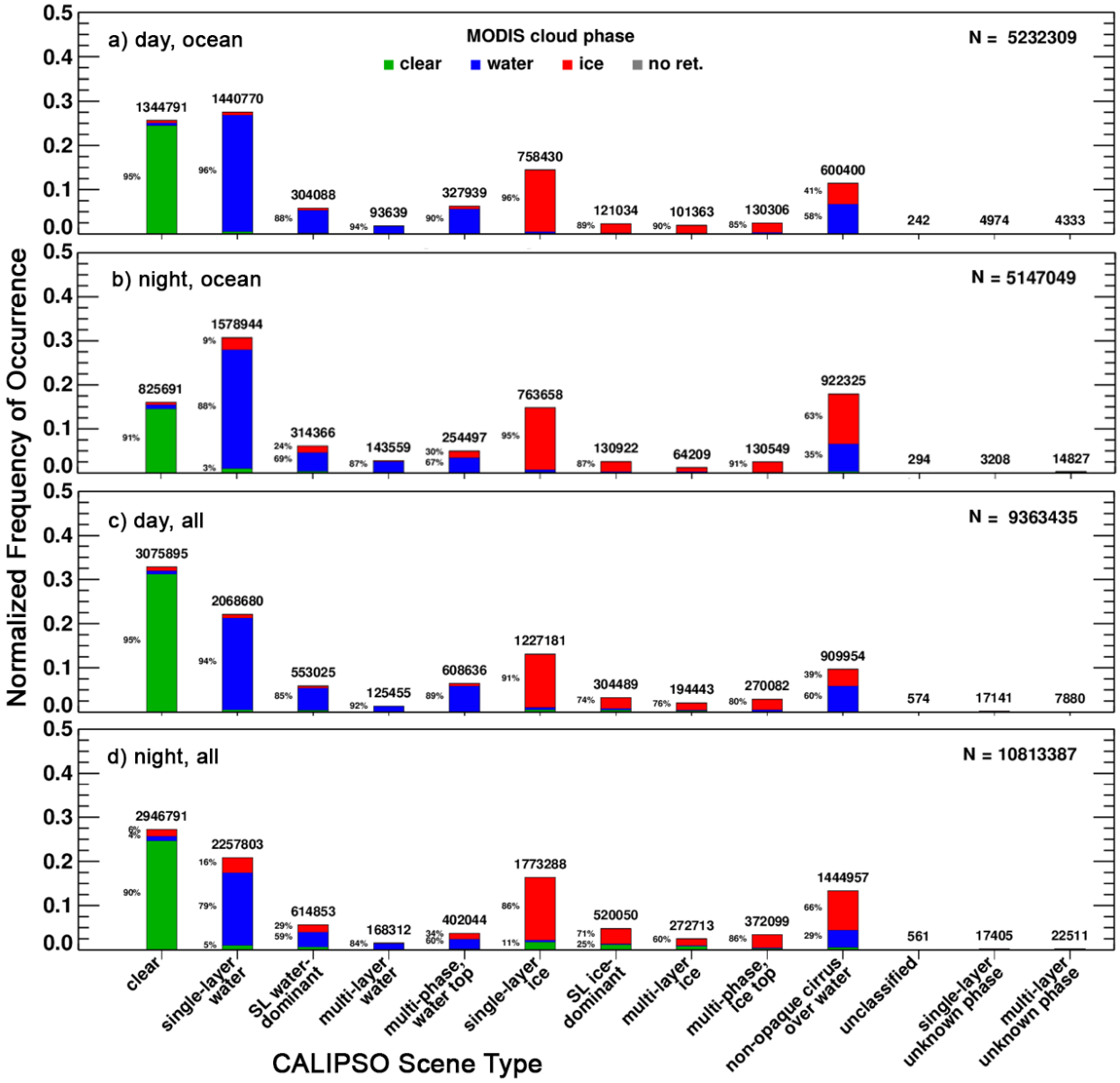

Fig. 19. Aqua MODIS Ed4 pixel scene classification as a function of CALIOP scene classification for overcast or totally clear CALIOP and MODIS pixels, JAJO 2015-16. Top: Ice-free ocean, Bottom: All surface types.

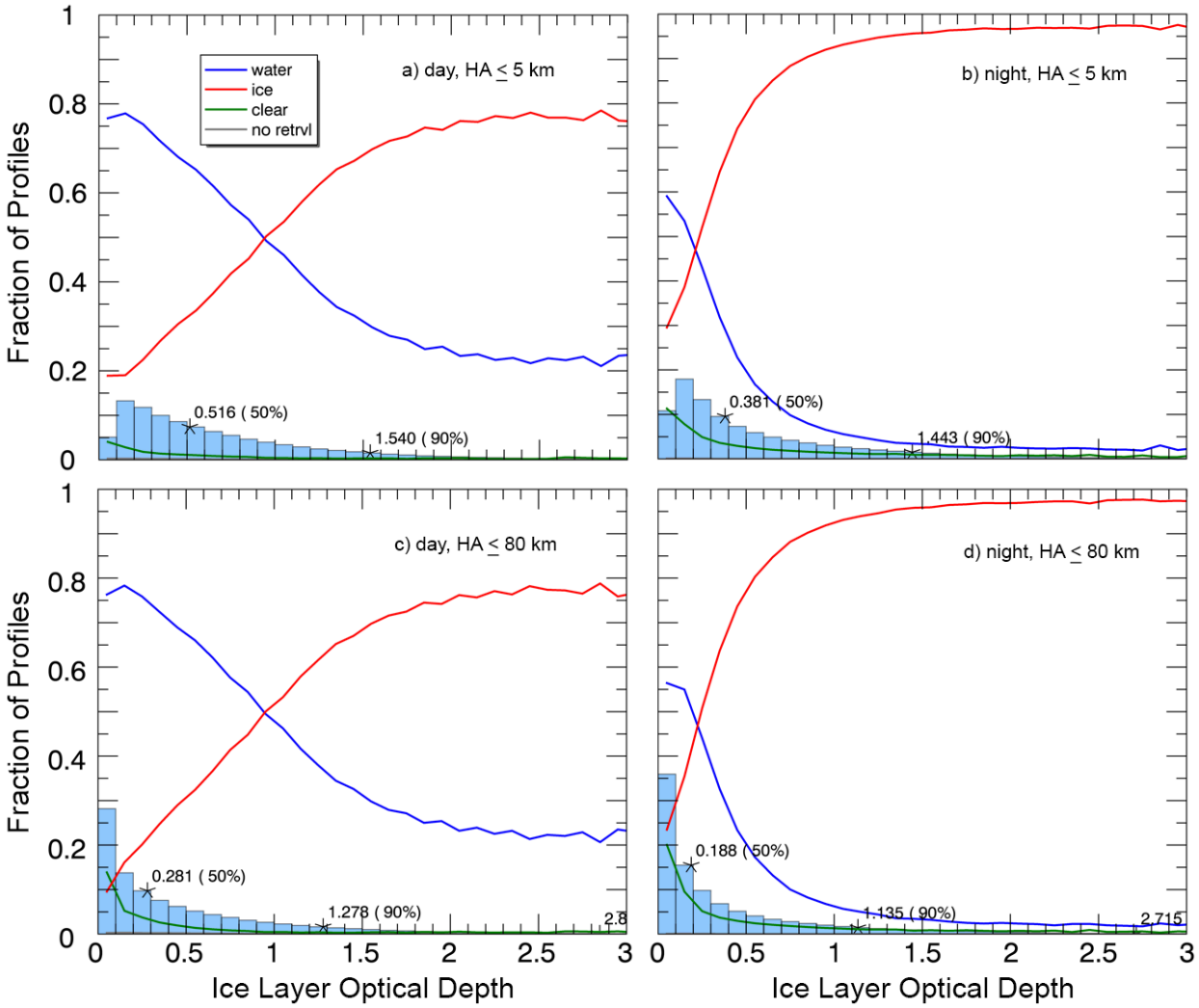

Fig. 20. Probability distribution of multilayer, semitransparent ice-over-water cloud pixels as identified by CALIOP during JAJO, 2015-16, Top: HA ≤ 80 km, Bottom: HA ≤ 5 km, Left: day, Right: night. Fraction of multilayer, CALIOP ice-over-water cloud pixels identified as clear (green), ice (red), or liquid water (blue) clouds by Ed4 within each 0.1 optical depth interval. Blue bars denote fraction of profiles having a CALIOP ice-layer optical depth within the 0.1 range.

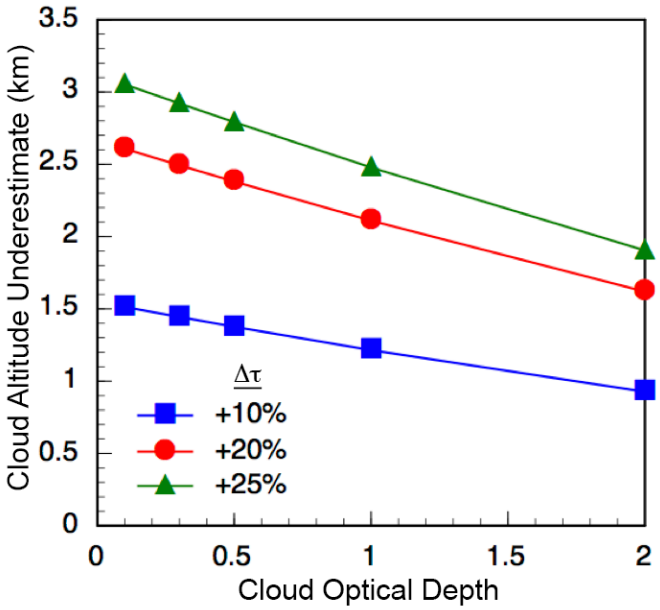

Fig. 21. Underestimate in cloud effective height as function of cloud optical depth and errors in optical depth based on cloud with effective temperature at 220 K and surface temperature of 290 K. Optical depth errors, $\Delta\tau$ (colored symbols) shown as percentage of true value of COD.

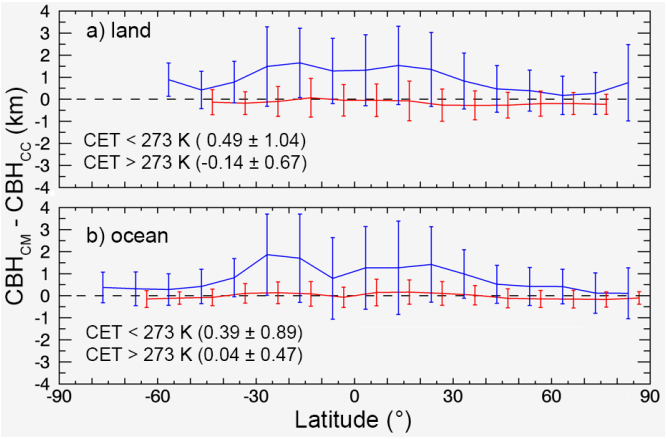

Fig. 22. Mean daytime differences between CERES Ed4 Aqua MODIS and CALIPSO-CloudSat single-layer liquid cloud base heights for snow/ice free regions as a function of latitude, JAJO 2010. All data used require that the nearest 2 pixels on each side of the CC pixel are liquid clouds.

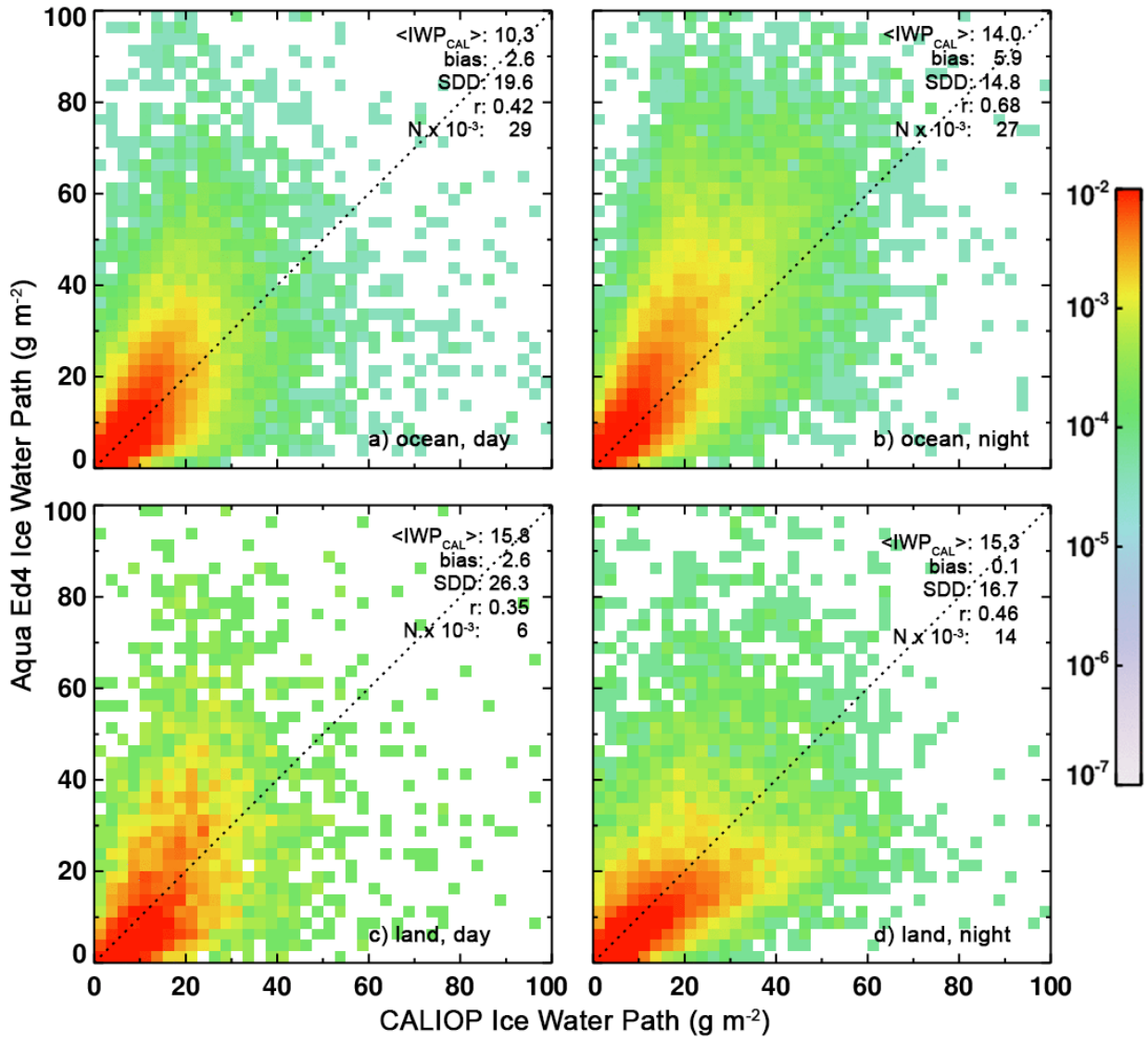

Fig. 23. Same as Fig. 17, except for July 2013 using Version 3 CALIPSO Layers Product.